Effects of Varying Degrees of Musical Complexity on Human Psychophysiological Measures

A Senior Honors Thesis

Presented in Partial Fulfillment of the Requirements for Graduation with Distinction in Psychology in the Undergraduate Colleges of The Ohio State University

By

Benjamin O. Turner

The Ohio State University

June, 2007

Project Advisor: Dr. Julian Thayer, Department of Psychology

Abstract

Recent studies have shown that listening to music can effect changes in heart rate variability (HRV) and respiratory rate. One dimension of music that has not been studied psychophysiologically is complexity. The present study examined psychophysiological responses to variations in musical complexity. Measures included high- and low-frequency bands of HRV, stroke volume, cardiac output, and the systolic time intervals (pre-ejection period and left-ventricular ejection time). Participants listened to ten two-minute songs of high or low complexity while their physiological responses were recorded. The hypothesis that increased complexity would be associated with decreased parasympathetic nervous system activation was partially confirmed. This research has implications for the music industry, as well as on theories of the perception of stimulus complexity.

Effects of Varying Degrees of Musical Complexity on Human Psychophysiological Measures

Stimulus complexity is a little-researched topic in most areas of psychology, and where there is research on it, such research tends to be isolated and domain-specific. However, there is no reason complexity should be ignored; everything from interpersonal interactions, to career choice, to artistic preference is influenced by a person's tolerance for complexity. The present experiment uses music as the stimulus in an examination of complexity's impact. The domain of music is one which does have a relatively rich foundation in the study of complexity, but we propose to examine the issue from a novel point of view, namely that of psychophysiology.

In many ways, this research is a continuation of the tradition established by Daniel Berlyne , virtually the founder of the area of experimental aesthetics. It was his belief that curiosity is a human drive, like any other biological drive. In his research, he often spoke of the novelty, interestingness, and complexity of stimuli. Further, he established a biological foundation for these constructs, based on the concept of arousal: he posited the existence of an inverted-U-shaped relationship between complexity and liking, based on an individual's search for his or her optimal level of arousal such that suboptimal levels lead the individual to seek stimulation while superoptimal levels lead the individual to reduce stimulation (Berlyne, 1970). As such, he occasionally employed psychophysiological approaches in his research, e.g., Berlyne, Craw, Salapatek, and Lewis (1963), in which the experimenters measured galvanic skin responses (GSR) as a measure of orienting response, with the finding that "There is thus some indication of a greater incidence of GSRs with more complex or incongruous visual patterns," but that these effects "are certainly not comparable with the pronounced effects that these same variables have on exploratory behavior" (p. 567).

However, since Berlyne's death in 1976, there has been very little work examining the interaction between complexity and psychophysiology, and virtually none (Birbaumer, Lutzenberger, Rau, Mayer-Kress, & Braun, 1996, offers a rare exception) in the field of music perception. This is not a symptom of a lack of psychophysiological research in music perception – numerous studies examining the psychophysiology of emotion (e.g., Nyklícek, Thayer, & Van Doornen, 1997; Etzel, Johnsen, Dickerson, Tranel, & Adolphs, 2006) and arousal responses to music (e.g., Iwanaga, Kobayashi, & Kawasaki, 2005; Iwanaga & Moroki, 1999) have been carried out recently. Nor is it due to a lack of study of complexity in music perception – papers examining the role of complexity in liking, most admittedly inspired by Berlyne, are plentiful (e.g., Burke & Gridley, 1990; North & Hargreaves, 1995; Orr & Ohlsson, 2005), as are papers studying theoretical approaches to quantifying complexity in music (e.g., Pressing, 1998; Streich, 2005). Rather, it seems that each of the fields is essentially unaware of the other's existence.

A short review of the research in both veins will help to set the stage for the experiment to be described below. To begin with, let us examine the music-theoretic approach. The topic of complexity is certainly not a new one in music theory, but it has only just recently begun to see serious treatment by more psychologically-oriented theorists. A brief survey of the more recent work reveals that much of it lies in the domain of information theory. In 1998, Pressing raised some questions about the classification of musical complexity, identifying three ways of measuring it: hierarchical complexity, which has to do with the structure of the music; dynamic complexity, which focuses on time behavior and change; and information-based complexity, which takes its cue either from Shannon (1948; cited in Pressing, 1998) and considers entropy (a psychologically implausible measure) , or from computer science, where it relates to programming cost (i.e., what is the shortest program that could be written that would be able to reproduce the stimulus?). Eerola and North (2000) posited an expectancy-based model of melodic complexity, which was highly able to predict participants' complexity ratings in an experiment they conducted. Also in 2000, Shmulevich and Povel proposed a novel measure of temporal pattern (i.e., rhythmic) complexity. They discuss the types of complexity listed in Pressing (1998), and also examine existing measures of rhythmic complexity, the Tanguiane (T-) measure and the Lempel-Ziv (LZ-) measure, and find both to be inadequate when looked at from a perceptual standpoint. Their measure is based on the work of Povel and Essens (1985; cited in Shmulevich & Povel, 2000) and assumes the existence of an internal clock. It achieves a correlation of $r = 0.75$ with actual judgments of complexity given by participants, as compared with $r = 0.02$ and $r = 0.15$ for the T- and LZ-measures, respectively. Finally, Streich (2005) attempted to create a program that could automatically characterize musical complexity, as part of a music tracking and organizing program. He reviewed much of the work done in musical complexity over the past several decades, and wrote a program that uses a "danceability" measure to achieve moderate success at differentiating songs based on their complexity.

In addition to these theoretical articles, there is a literature in the tradition of Berlyne, examining the relationship between complexity and liking, generally as a function of some other factor. Berlyne hypothesized an inverted-U-shaped curvilinear relationship between complexity and liking, such that liking increases up to a point with increasing complexity, then starts to decrease as the stimulus becomes too complex. Steck and Matchotka (1975) found support for this relationship, but also found that it was context-dependent rather than absolute. Burke and Gridley (1990) examined how musical preference was affected by both stimulus complexity and listeners' sophistication, and found inverted-U-shaped curves for liking as a function of complexity, shifted up for sophisticated listeners. North and Hargreaves (1995) did a similar experiment, except that they looked at the role of familiarity, rather than sophistication, and found support for Berlyne's hypothesis as well as a positive relationship between familiarity and liking. The potential interaction between familiarity and complexity, that increased familiarity generally leads to decreased perceived complexity, was discussed. Finally, Orr and Ohlsson (2005) examined complexity and liking as a function of expertise, and found no consistent relationship, leading them to posit that "musical expertise dissolves the relationship between liking and complexity" (Orr & Ohlsson, 2005, p. 583).

Multiple studies have also been carried out that employ complexity manipulations in music and examine the effects on some other aspect or task. Kiger (1989) found that performance on a reading comprehension task was as low in a high information-load condition, operationalized as being "a dissonant, rhythmically varied and highly dynamic piece," as it was in a silent condition (p. 532). However, performance was not impeded by a low information-load condition, operationalized as being "a highly repetitive synthesizer piece with a narrow tonal range" (p. 532). This led Kiger to posit the role of arousal, as a function of complexity, in task performance. A decade later, North and Hargreaves (1999) examined music complexity as it related to waiting time. Participants were left alone, believing the experimenter would return shortly. Although there was a ceiling effect (the experimenter always did actually return after 20 minutes), they concluded that participants waited the least amount of time during a no-music condition, and that there were no differences in waiting time among music conditions of different complexity. Finally, in an examination of how complexity influences success on the Billboard charts, Parry (2002) found a positive correlation between complexity and overall chart performance.

One more study needs to be mentioned before moving on to the field of psychophysiology, a study that examined how complexity is perceived in short periods of time. Scheirer, Watson, and Vercoe (2000) examined judgments of complexity for 5-second clips of preexisting music, and found that "the perceived complexity of a musical signal is an important surface feature of music." In other words, even to naïve listeners hearing only five seconds of music, complexity is not an abstract, theoretical construct but rather a concrete, important element of music. This primacy of complexity as a facet of music provides support for the efficacy of our dependent manipulation.

In our study, the focus of psychophysiological investigation is on cardiac function, specifically those measures obtained by electrocardiography (ECG) and impedance cardiography (ICG). From a more purely physiological stance, the autonomic nervous system is divided into the sympathetic and parasympathetic branches, which act in accentuated antagonism on the heart (Levy, 1971; Uijtdehaage & Thayer, 2000). The relative degree of control exerted by either branch can be estimated using a combination of variables, including heart rate, heart rate variability and the systolic time intervals (left-ventricular ejection time and pre-ejection period) (Saul, 1990; Malliani, 1999). Among other things, some of the interactions between the parasympathetic and sympathetic branches have been attributed to cognitive load, a hypothesis particularly relevant to a study of the effects of complexity (Allen & Crowell, 1989; Tanaka, Sawada, & Fujii, 1994).

Music has been examined with psychophysiological methods since Kate Hevner in the 1930's, and probably before that. However, in all that time, it does not appear that complexity was ever carefully examined. Nearly all of the studies in recent years have focused either on how arousing the music is, or on what emotions it induces in the listener. A brief summary of some of these studies follows, beginning with those looking at music as a type of sound, and then moving on to those considering music in its own right.

Bartlett (1996) presents a thorough review of existing research on psychophysiology and sound, examining responses among all the major physiological systems. Yanagihashi, Ohira, Kimura, and Fujiwara (1997) examined the psychophysiological effects of sound, including a music condition, and found that mechanical sounds inhibit the parasympathetic nervous system significantly more than do either bird twitters or music. Gomez and Danuser (2004) presented participants with 16 environmental noises and 16 musical fragments, and found that breathing accelerated with increases in both valence (positive vs. negative emotions) and arousal (high vs. low energy), as rated by the participants. Additionally, skin conductance levels increased with arousal ratings for music but not noise, while the opposite was true for mean heart rate. This set of results suggests that breathing measures and heart rate may be good measures of emotion induction using music.

Thayer (1986) was able to differentiate affective responses to music along the two dimensions of valence and arousal using electromyography and electroencephalography. Nyklícek, Thayer, and Van Doornen (1997) examined the cardiorespiratory differentiation of musically-induced emotions, and found that when all the dependent variables were considered together, a discriminant analysis allowed clear physiological differentiation of four different emotions. That same year, Iwanaga and Tsukamoto (1997) examined the influence of sedative and excitative music on physiological relaxation, taking heart rate variability as their primary dependent measure, and found that the music acted on the parasympathetic but not the sympathetic nervous system. Iwanaga and Moroki (1999) did a similar study in which they found that whether music was excitative or sedative, but not a listener's music preference, influenced cardiorespiratory measures. Iwanaga, Kobayashi, and Kawasaki (2005) found some effects indicating that parasympathetic deactivation became less pronounced with repeated exposure to the same piece of music. In a cross-modal study, Baumgartner, Esslen, and Jäncke (2006) found that music markedly enhanced emotional experiences evoked by affective pictures, as measured by EEG, heart rate, skin conductance, respiration, and temperature. Nater, Abbruzzese, Krebs, and Ehlert (2006) found sex differences in emotional and psychophysiological responses to musical stimuli, such that women tended to show hypersensitivity to aversive musical stimuli, as reflected in heart rate, electrodermal activity, skin temperature, and pharmacological measures. Also in 2006, Etzel, Johnsen, Dickerson, Tranel, and Adolphs found little variation between different moods as measured by cardiovascular and respiratory patterns, but concluded that tempo differences may have led to these inconsistent results. Supporting this interpretation, Bernardi, Porta, and Sleight (2006) reported that music induces an arousal effect, as measured by cardiovascular, cerebrovascular, and respiratory changes, related primarily to tempo.

Another psychophysiological study must be mentioned, which looked at music preference as a function of sensation seeking. Nater, Krebs, and Ehlert (2005) found that although participants high in sensation seeking (Zuckerman, 1994) indicated that they felt less activation after arousing music than did low sensation seekers, no differences were found between their physiological measures and those of low sensation seekers. This finding is interesting, because the physiological and psychometric measures disagree. A short review of sensation seeking and related concepts is necessary, because they will be considered in the present experiment.

Sensation seeking is a measure created by Zuckerman, "defined by the need for varied, novel, and complex sensations and experiences and the willingness to take physical and social risks for the sake of such experience" (SSS; Zuckerman, 1994, p. 2). It has been tied both to psychophysiology (Zuckerman, 1990) and musical preferences (Litle & Zuckerman, 1986). Other measures which might prove relevant in a person's preference for, or reactivity to, complex stimuli include Cacioppo and Petty's Need for Cognition, which "refers to an individual's tendency to engage in and enjoy effortful cognitive endeavors" (NCS; Cacioppo, Petty, & Kao, 1984, p. 306), and the Perceptual Curiosity Scale, supposed to measure perceptual curiosity, which is "evoked by complex or ambiguous patterns of sensory stimulation" and motivates "behaviors such as visual inspection in order to acquire new information" (PCS; Collins, Litman, & Spielberger, 2004, p. 1128).

A person's scores on each of these scales could be expected to play a moderating role in responses to complexity. However, it is not clear which direction the interaction could be expected to operate. The only scale that has been examined psychophysiologically is the SSS, where Zuckerman reports that "high sensation seekers tend to give stronger physiological orienting responses than lows to novel stimuli of moderate intensity, particularly when such stimuli are of specific interest. Lows tend to show defensive responses as defined by heart-rate acceleration" (Zuckerman, 1990, p. 313). However, he also states in the conclusion, "The evidence presented in this article associating sensation seeking with [physiological responses] is not always consistent and seems to depend on the specific parameters of the stimuli…" (Zuckerman, 1990, p. 339). The author's ultimate interpretation is that high and low sensation seekers may have different evolved biological strategies to deal with ambiguous stimuli, such that high sensation seekers investigate moderate stimuli and tolerate high-intensity stimuli, where low sensation seekers ignore or avoid moderate stimuli and shut down in response to high-intensity stimuli. Therefore, it may be the case that high sensation seekers will be more reactive to complex stimuli, but the mixed evidence showing defensive responses in low sensation seekers makes a specific prediction difficult.

Similarly, conjectures could be made about the role of need for cognition and perceptual curiosity in moderating a subject's physiological response to complex stimuli, but there is little research on which to draw in such speculation. The significant positive correlations between the PCS and selected subscales of the SSS (Collins, Litman, & Spielberger, 2004) suggest that the PCS may interact with physiological responses in a manner similar to the SSS. Even less evidence exists for the NCS; while the NCS was utile in determining participants' attitudes towards simple and complex versions of a cognitive task, such that participants high in need for cognition preferred the complex to the simple task, and participants low vice versa, the physiological implications of this effect are not known (Cacioppo & Petty, 1982).

As mentioned above, there has only been one study to recently study psychophysiology and musical complexity, namely Birbaumer, Lutzenberger, Rau, Mayer-Kress, and Braun (1996). The researchers varied the complexity on the rhythmic dimension, the melodic dimension, or both dimensions by having either periodic sequences, sequences with intermittent chaos, or quasi-random sequences. The findings of this study suggested that weakly chaotic music entrained less complex brain wave oscillations at the prefrontal cortex than did either strongly chaotic or periodic music; subjective ratings of complexity did not align with the changes in brain wave complexity.

There are two hypotheses under investigation in the present study. The first, examining the influence of musical complexity on psychophysiology, is that music of higher complexity will induce a greater parasympathetic withdrawal. As discussed above, the deactivation of the parasympathetic branch of the autonomic nervous system has been hypothesized to reflect increased cognitive load, an interpretation that would make sense if participants were engaged in interpreting higher-information (i.e., more complex) stimuli. Because of results in previous studies that found this parasympathetic decrease with no attendant sympathetic increase (Iwanaga & Tsukumoto, 1997), we will posit the same pattern of results in the present study.

The second hypothesis is that participants' scores on the SSS, the PCS, or the NCS will have a moderating role on this physiological difference. As discussed above, there is mixed evidence with regards to the interaction between sensation seeking and psychophysiology, and no evidence at all on either of the other two scales and their psychophysiological bases or effects. Therefore, this hypothesis is necessarily bi-directional for each scale.

Method

Participants

Thirty-eight undergraduates (18 females and 20 males) between the ages of 18 and 25 (mean age = 19.1, standard deviation = 1.3; one participant did not give her age) enrolled for participation in the experiment in partial fulfillment of a research experience option for an introductory course in psychology. There was no difference between the ages of the two genders ($t(35) = .401$, $p = .680$). Participation was limited to those self-identifying as non-smokers, and participants were asked to refrain from alcohol and caffeine for the 24-hour period prior to their participation in the experiment. Data were discarded or never collected from a number of participants because of software malfunction (n = 1), programming errors (n = 1), and hardware malfunction (n = 6).

Musical stimuli

Stimuli were existing pieces of music, chosen while taking into consideration the potential confounds discussed below. In order to have as ecologically valid a set of stimuli as possible, five works were chosen, three in the major mode and two in minor, and tempi and ranges were all musically valid (see Table 1 for the list of works). To achieve the desired complexity manipulation, each of these melodies had a second, more ornamented version (see Figure 1 for an example of ornamentation and a lengthier discussion of the ornamentation procedure). While we took our method of creating a complex version from the musicological idea of theme and variation, wherein certain aspects of a basic "theme" piece are expanded upon while other elements are left fundamentally unaltered, one can view our result as translating to differences in the number of notes per unit time. Summary data for each piece in both of its versions are presented in Table 1.

In this manner, two distinct levels of complexity were presented without diminishing the ecological validity of the stimuli. All stimuli were the same intensity, and none of the ornaments significantly increased the overall range of any of the melodies. Other confounds intentionally avoided include dissonance and tempo; while the five main melodies could vary slightly from one another on either of these measures, none of the manipulated versions differed from its original in either of these regards. The stimuli were created using MakeMusic, Inc.'s Finale® 2003 and exported as MIDIs, which were then converted to 16-bit, 44kHz mono WAV files using COWON America, Inc.'s jetAudio converter. These WAVs were finally converted to stereo and reduced to 22kHz in Audacity, and were split in half in order to run properly in the experiment software. Unfortunately, this splitting created a small clip in the middle of the song, but several participants were asked and reported being unaware of it. All stimuli were presented over Sennheiser HD 650 headphones, played using Psychology Software Tools, Inc.'s E-Studio v1.2 which was run on a Dell OptiPlex GX620 in the control room and fed into the subject room over a Mackie 1642-VLZ Pro mixer.

Psychometric variables

Multiple questionnaires were administered to participants at various stages of the experiment. First, because anxiety is known to impact measures of cardiovascular function, the six-item short-form of the Spielberger State-Trait Anxiety Inventory (STAI) was administered (Marteau & Bekker, 1992). To examine the impact of musical taste or expertise, an abbreviated version of the Musical Preference Scale (MPS, Litle & Zuckerman, 1986; see Appendix A) was employed. Further, to assess the role of personality factors in mediating physiological responses to complexity, the following scales were administered to all participants: the short form of the NCS (Cacioppo, Petty, & Kao, 1984); the PCS (Collins, Litman, & Spielberger, 2004); and the SSS (Zuckerman, 1994). Finally, as a validation measure for the complexity manipulation, participants were asked to judge each stimulus on a music evaluation questionnaire (MEQ, Appendix B). All questions were presented in white lettering on black background on the screen of a Hitachi Ultravision 42HDS69 television, at a distance of approximately 6'6" in front of the participant, using E-Studio. All answers were made by the participant using the Microsoft Bluetooth® wireless keyboard and mouse in the subject room and collected by E-Studio.

Physiological variables

All physiological measures were taken using MindWare Technologies Ltd.'s MW2000 ICG amplifier and leads, connected to BIOPAC Systems, Inc.'s EL503 electrodes, recorded using MindWare Technologies Ltd.'s VideoACQ v1.2.1 software and compiled and filtered using MindWare's Heart Rate Variability and Impedance Cardiography software programs.

These programs automatically detect muscle artifacts in the EKG signal, and the user can manually choose to delete spurious data points. They also calculate the very low frequency (VLF), low frequency (LF) and high frequency (HF) bands of the heart rate variability (HRV), the low frequency/high frequency ratio (LH), heart rate (HR), respiration rate (RR), inter-beat interval (IBI), left ventricular ejection time (LVET), stroke volume (SV), cardiac output (CO), and pre-ejection period (PEP) (see below for further discussion).

Experimental design

The primary independent variable of interest in this experiment was complexity. All participants experienced both levels of the independent variable, presented in a pseudo-random order (order was constrained by the rule that the same melody could not occur twice in a row). The subject variables of sensation seeking, perceptual curiosity, need for cognition, and state anxiety were also measured for all participants, but no attempt was made to select for participants on the basis of any of these measures. The dependent variables were an array of psychophysiological measures, *a priori* focusing on IBI, LF, HF and LVET, which are all thought to be good indicators of either sympathetic or parasympathetic nervous system activity. The inclusion of all measures will allow for a distinction between diminished parasympathetic activity and increased sympathetic activity. In total, there was a 2 (complexity – simple/complex) x 5 (song – 1/2/3/4/5) x 2 (period – rest/trial) within-subjects design, with 13 possible between-subjects variables (gender, scores on the SSS, NCS, PCS, MPS, and all subscales, initial STAI score and STAI score change from initial to final, labeled ΔSTAI), ten dependent physiological measures, and answers to the six MEQ questions following each of the songs.

Procedure

After providing their informed consent, participants were seated in one of the subject rooms in a comfortable chair, facing the television screen. The experimenter, who was always gender-matched with the participant, explained that the procedure involved the placement of electrodes on the participant's torso and neck. After the participant verbally indicated that he or she was comfortable with the procedure, the experimenter proceeded to attach the electrodes. Once all the sensors were wired to the amplifier, participants were read a script briefly describing the experiment – they were specifically made aware that they were free to leave at any time without penalty, and that they were being monitored but not recorded via a small network camera located above the television screen. After the instructions were finished, participants were asked if they had any questions, and then instructed to put on the headphones. After filling out the STAI, participants rested quietly for 220 seconds, with a warning tone and visual message one minute before the onset of the music and another warning tone and accompanying visual countdown ten seconds before the onset of the music, in order to minimize any potential startle response. Each stimulus was between 90 and 120 second in length (see Table 1), and was followed by a rest period of at least 110 seconds. The total interval duration between songs was determined by the participant: at the end of each song the participant was instructed to hit a key to continue on to the MEQ, and E-Studio would pause until he or she hit a key, both at that initial instruction and also during each MEQ question. The average span of time between two consecutive songs across participants was 152.82 seconds (sd = 13.26 seconds), with a minimum rest across all participants of 135.24 seconds and a maximum of 255.89 seconds. The mean total duration for presentation of all ten stimuli, with breaks, was 42:43 (sd = 1:13), and physiological recording was continuous throughout the experiment. Following completion of the tenth MEQ and subsequent rest period, participants completed the STAI a second time, the SSS, the NCS, the PCS, and the MPS, after which they were disconnected from the physiological equipment, debriefed, thanked for their participation and dismissed.

Data analysis

Analyses were performed for the 30 participants whose data were collected successfully (16 males, 14 females; mean age = 19.2; age difference between genders $t(27) = .187$, $p = .853$). The physiological variables were always computed over a 90-second span constrained to end with the end of each period. Using MindWare's HRV software, HR and IBI means were calculated first in ten-second intervals, and then across all nine intervals in the period. Voltage full scale, sampling frequency, and A/D resolution were all constants from the input file; their values were 10.00 volts, 1000.00 Hz and 16-bit respectively. Dz/dt was used as the respiration signal. For the HRV analysis, the IBI series was subjected to a Hamming window with the following ranges: $0.003 \leq \text{VLF} \leq 0.040$; $0.040 \leq \text{LF} \leq 0.150$; $0.150 < \text{HF/RSA} \leq 0.400$ (all units Hz). LF and HF were natural-log transformed prior to analysis, though the LF/HF ratio was computed on the untransformed values. For the impedance measures, MindWare's IMP software had the following settings: Zo calibration = 0.10 Volts/Ohm, dZ/dt calibration = 0.50 Volts/Ohm per sec, Rho = 135, Ensemble Window Max = 550ms, Ensemble Window Min = 100ms, K constant = 35, and percent dZ/dt peak = 56%. The minimum value K-R Q-point calculation method was used, along with the percent dZ/dt time B-point calculation method. Distances between front and back impedance electrodes were measured for each participant and entered as constants in the program as well.

To clean the data, MindWare's automatic outlier-detection algorithm was allowed to highlight potential noise in the data. In each period, these highlighted points were manually examined and, if noise did exist, removed and replaced as appropriate. In some instances, recording was seriously impaired for the duration of an nearly an entire period. In those cases, a note was made of the problem, and if there was no way of getting means from that period, the participant was excluded from analysis for all variables with significant loss.

Results

Figure 2A shows mean ratings of complexity for each song, plotted against mean number of notes per unit time (calculated as the average number of notes per measure adjusted for tempo) on the abscissa, and Figure 2B shows the same relationship for the raw ratings of complexity. As shown on the graphs, participants were highly perceptive of the manipulation, with $r = 0.99$ for the averaged values and 0.84 for the raw values. No other aspect of the songs that we examined (i.e., those listed in Table 1) predicted complexity ratings so successfully.

However, looking at the correlations between responses (averaged across participants) on each of the six MEQ questions, it is obvious that complexity is confounded with each of the other dimensions that we investigated. The correlations between complexity and the other five aspects measured by the MEQ were: excitingness, $r = 0.99$; relaxingness, $r = -0.67$; happiness, $r = 0.91$; sadness, $r = -0.87$; and liking, $r = 0.84$. Four of these (excitingness, happiness, sadness, and liking) were significant at the $p < .05$ level after correction for multiple observations. A method for dealing with these confounds is discussed later.

In looking at the physiological data, repeated measures GLM tests were used with a 2 (complexity) x 5 (song) x 2 (period) design. Thus, the effect of interest was a complexity x period interaction, such that complex and simple stimuli would show differential changes from their respective baselines. Because order was pseudo-random, there should of course have been no differences between baselines in simple and complex conditions; this assumption was not always upheld, and individual deviations from this assumption are discussed below.

In addition to this 2 x 5 x 2 analysis, the entire range of subject variables could be entered into the model as between-subjects factors, or as covariates if appropriate. The subject variables to be examined include gender, along with scores on: the SSS and its four subscales; the NCS; the PCS and both its subscales; the MPS; the initial STAI; and ΔSTAI. Means, standard deviations, and other descriptives of each subject variable are presented in Table 2.

A repeated-measures GLM was applied to the data with all possible physiological variables, after removal of those participants whose data were unreliable on one or more of the variables. However, even with an *n* of only 22 (12 females, 10 males), the complexity x period interaction in the model including gender was significant for four of the variables: LVET ($F(1,20) = 6.56$, $p = .019$), RR ($F(1,20) = 4.64$, $p = .044$), SV ($F(1,20) = 7.65$, $p = .012$), and CO ($F(1,20) = 9.48$, $p = .006$). Further, the complexity x period x gender interaction in the same model was significant for two of those variables, SV ($F(1,20) = 8.18$, $p = .010$) and CO ($F(1,20) = 9.87$, $p = .005$). Means, standard errors, and confidence intervals for all variables of interest are presented (separated by gender) in Tables 3A-3H.

In looking at the variables individually, there were five for which the complexity x period interaction was significant or neared significance. For each variable, any participant whose data were identified as extreme values (defined by SPSS v.15 as values more than three box lengths from the end of the box in a boxplot of the data for all participants) on multiple trials was removed from analysis for that variable. For CO, three participants were removed from analysis. As a further check, each participant was removed in a rotated fashion from the full set of 30 participants; for only one removal was the *p* value less than .05. When the same procedure was carried out with the remaining 29 participants, p was less than .02 in every case. This pattern of results suggests that this one participant, who was one of the three identified as having extreme values, was artificially driving p above .05. After all three participants were removed, the complexity x period interaction was significant, $F(1,25) = 6.55$, $p = .017$. Further, there was a significant complexity x period x gender interaction, $F(1,25) = 6.86$, $p = .015$. Individual t-tests, Bonferonni corrected, revealed the only significant effect to be for males in the simple condition ($t(12) = 2.79$, $p = .008$), where there is a significant increase from rest to trial. This is shown graphically in Figure 3.

Because of its relationship to CO (discussed below), the same situation exists with SV. There was a significant complexity x period interaction, $F(1,25) = 4.39$, $p = .046$, and a significant complexity x period x gender interaction, $F(1,25) = 5.32$, $p = .030$. Individual t-tests again reveal a significant change from rest to trial for males in the simple condition ($t(12) = 3.48$, $p = .002$), shown graphically in Figure 4.

The interaction for HF approached significance in an individual multivariate test, $F(1,26) = 3.21$, $p = .085$. This result came after removal of two participants whose data were unreliable; as before, each participant was removed one at a time in rotation to determine the robustness of the result. This time, there was a 3:1 ratio of participants whose removal caused p to rise above .1 to those whose removal caused p to fall below .05; however, never was p above .15, which suggests that the result is trending towards significance. This conclusion is supported by a single significant effect in individual corrected t-tests, where the change from baseline to trial for males in the complex condition was significant ($t(13) = 2.90$, $p = .006$). Visual inspection of the graph shows this relationship to be a decrease in HF from baseline to trial for males, with no change in the simple condition for males or in either condition for females (see Figure 5).

For RR, the complexity x period interaction likewise approached significance, $F(1,28) = 3.99$, $p = .055$ (no participants were removed from analysis). Individual t-tests revealed a significant change for both males and females from baseline to trial in the complex condition ($t(15) = 2.65$, $p = .009$; $t(13) = 2.99$, $p = .005$) (see Figure 6). Because of the interaction between RR and HF, subsequent tests were carried out to determine the degree to which the differences in HF were driven by changes in RR. The first line of evidence to suggest that HF is not driven solely by RR is that for HF, only males showed a significant difference, in the complex condition ($t(13) = 2.90$, $p = .006$; females, $t(13) = -0.298$, $p = .615$), while for RR both genders showed a significant difference (males, $t(15) = 2.65$, $p = .009$; females, $t(13) = 2.99$, $p = .005$), all from baseline to trial in the complex condition. The second line of defense is that while the difference between the change in simple and complex conditions has a Cohen's *d* of .21 for RR, the absolute difference was less than 0.6 breaths per minute, or 3.6% of the total average RR. The final line of evidence that HF is independent of RR comes from the correlations between the two: overall, $r^2 = .026$; for males, $r^2 = .066$; and for females, $r^2 = .209$. Thus, RR explains very little of the variability in HF overall and for males, although for females it is more of a driving force.

The final variable for which there was a significant complexity x period interaction was LVET, $F(1,28) = 4.46$, $p = .044$ (no participants were removed from analysis). Although there was no significant complexity x period x gender interaction ($F(1,28) = 1.87$, $p = .183$), individual t-tests revealed a significant change from baseline to trial only for males in the simple condition ($t(15) = 3.33$, $p = .002$). This variable actually shows an unusual pattern upon closer inspection: females show a decrease in both conditions (neither significant: simple, $t(13) = 1.25$, $p = .125$; complex, $t(13) = 1.57$, $p = .069$), while males show the aforementioned change, an increase, in the simple condition, and no change in the complex condition ($t(15) = -.62$, $p = .728$)

(see Figure 7). Taken together, these changes create a picture wherein LVET increases slightly (but nonsignificantly) to simple stimuli and decreases slightly more (but still nonsignificantly) to complex stimuli. Of course, this is a false outcome, the result of glossing over significant differences between the reactions of males and females.

Before moving on to the psychometric interactions, there were significant effects of period without any significant complexity x period interaction for three other variables, meaning that although complexity did not modulate the change, music caused a reaction. For LF, the main effect of period was highly significant, $F(1,27) = 21.80$, $p = .001$ (one participant was removed from analysis). Individual t-tests revealed significant changes for females in the simple condition ($t(13) = 2.97$, $p = .005$) and for males in the complex condition ($t(14) = 3.47$, $p = .002$), with results nearing significance for females in the complex ($t(13) = 2.20$, $p = .023$) and males in the simple ($t(14) = 2.21$, $p = .021$). In all cases there was a decrease in LF from baseline to trial (see Figure 8). For LH, the main effect of period was again highly significant, $F(1,24) = 17.13$, $p = .001$. Individual t-tests revealed significant changes for females in the simple ($t(13) = 2.82$, $p = .007$) and complex ($t(13) = 4.00$, $p = .001$) conditions, with results nearing significance for males in the simple ($t(11) = 2.21$, $p = .025$) and complex ($t(11) = 1.75$, $p = .053$) conditions. Again in both cases, there was a decrease in LH from baseline to trial (see Figure 9). And finally, there was a significant main effect of period on IBI, $F(1,28) = 10.90$, $p = .003$. Individual t-tests revealed a significant change for females in the simple condition ($t(13) = 3.20$, $p = .003$) and for males in the complex condition ($t(15) = 2.56$, $p = .011$), with results nearing significance for females in the complex condition ($t(13) = 1.96$, $p = .036$). In all cases, there was an increase in IBI from baseline (see Figure 10).

In considering the influence of subject variables, again, a repeated-measures GLM was applied to all the variables. Each variable was entered first as a covariate, and then as a between-subjects factor with participants scoring in the middle third removed. In this manner, all possible interactions were identified for further investigation. Each subject variable that showed a significant complexity x period x *variable* interaction follows, given with its significant physiological variables: SS – LVET, SV, PEP; SS-ES – LVET, SV, CO; SS-DIS – LVET, SV, CO, PEP; SS-BS – IBI; NC – PEP; STAI – LVET, SV, CO; ΔSTAI – PEP. Each subject variable-physiological variable pair is tested more specifically below. It must also be noted that there were three subject variables that showed a significant difference between males and females: SS ($t(28) = -1.75$, $p = .091$), TAS ($t(28) = -2.17$, $p = .038$) and STAI ($t(28) = -3.03$, $p = .005$). As before, all means, standard errors and confidence intervals are presented in Tables 4A-4L.

For SS, the variables LVET, SV, and PEP were identified as being of interest. LVET showed a significant complexity x period x SS interaction when SS was entered as a covariate ($F(1,27) = 6.98$, $p = .014$), and a near-significant interaction when SS was split into low and high groups ($F(1,17) = 4.10$, $p = .059$), both with gender included as a between-subjects factor. T-tests (again, all t-tests were corrected for multiple comparisons) revealed no significant differences between low and high groups. In order to explore the covariate interaction, the following procedure was used: for each participant, the averages for each period in simple and complex conditions were computed (i.e., data collapsed across songs); differences between the trial and baseline were computed for both simple and complex conditions; differences between those values for simple and complex conditions were computed; and finally, a correlation was computed between these scores and participants' SS scores (see Figure 11 for the equation). The correlation was $r = -.495$, $p = .005$, which by itself does not give a full account of the interaction between complexity and SS, because it is necessary to determine which terms are negative to specify the direction of the effect.

In the case of LVET, change scores were generally negative in the complex condition (i.e., a decrease from baseline to trial) and positive in the simple condition (i.e., an increase from baseline to trial), leading to a negative complex - simple score. Of the other variables that showed a significant correlation between a psychometric and the double-difference score, SV and CO showed the same pattern of results, while PEP showed the reverse relationship (i.e., a decrease from baseline to trial in the simple condition and an increase in the complex). All significant correlations are shown graphically in Figures 12-18. The results of all the tests described in the above paragraph for all subject variable-physiological variable pairs are presented in Table 5. In every instance, the above procedures were used, with data appropriately cleaned to remove those participants identified as outliers.

As mentioned above, complexity was significantly confounded with most of the other aspects measured for each song. However, in looking at correlations between the each of these aspect scores (averaged across participants) and physiological responses (again averaged across participants), complexity provided the strongest correlation in most instances (shown in Table 6). Although the amount by which complexity is superior is not great, it nonetheless supports the idea of complexity as being the best overall way of looking at the stimuli.

Discussion

The implications of each of the significant interactions discussed above will now be addressed, looking first at the simpler complexity x period interactions, and then considering the more complicated complexity x period x *variable* interactions. The present pattern of results will then be addressed as they relate to existing research, and several avenues of possible future research will be suggested.

Of the five variables that showed a significant complexity x period interaction, HF power is dominated by the parasympathetic nervous system activity, while LVET is driven largely by the sympathetic branch. The interpretation of changes in SV, CO and RR are not as straightforward in terms of their locus of control in the autonomic nervous system. Adding to the difficulty in drawing clear conclusions from the pattern of results is the lack of any decrease in IBI (which would be consistent with a sympathetic nervous system increase) or change in LH (which has been described as a measure of the sympathovagal balance). However, there are clearly significant differences on several variables, so a tentative interpretation can be made.

As described above, the decrease in HF power is consistent with a decrease in parasympathetic activity. This result was seen only in males in the complex condition, suggesting that complex music initiates a withdrawal of the parasympathetic nervous system. A decrease in LF, or especially in LH, is generally interpreted as a decrease in sympathetic nervous system activation. This is because the sympathetic branch, being more slow-acting than the parasympathetic branch, has no effect on HRV in the HF band but has been posited to share influence with the parasympathetic branch in the LF band. However, depending on the location of the peak power in the LF band, one can infer whether the value derived for LF power is more dominated by vagal (peaks approaching 0.15, which was used as the upper bound for LF in the present experiment) or adrenergic (peaks lower than 0.15) inputs. Ultimately, IBI, LF, LH, LVET and HF (the latter two only for females) together paint a picture where sympathetic input decreases more markedly than sympathetic input across conditions, but this interpretation is extremely cautious because of the lack of any single clear, robust picture.

Without any measure of blood pressure, CO and SV are of limited use because total peripheral resistance cannot be obtained. Also, as mentioned above, they are related, specifically by the equation CO = SV x HR; therefore, the fact that SV and CO change in the complex condition by approximately equal amounts (in terms of total percent change) is consistent with the null change in IBI (60,000/HR). The results for LVET, as previously discussed, do not lend themselves to an immediate, obvious interpretation. The increase in LVET for males to simple stimuli is consistent with a decrease in sympathetic activation, while the (insignificant) decreases in LVET in both conditions for females is consistent with a slight increase in sympathetic control. However, this result is inconsistent with most of the other measures that suggest a decrease in sympathetic control, and might therefore be interpreted instead as a withdrawal of vagal control.

Difficulties in interpretation notwithstanding, it is evident from our analyses that our manipulations had an effect on several of the physiological measures examined in this experiment. It is further evident that sensation seeking or a similar trait moderated the degree of the effect that this manipulation had on autonomic function. Music of higher complexity initiated a withdrawal of vagal control, as demonstrated by the reduction in high frequency power, but this result was gender-specific such that it was observed only for males, and was probably accompanied by an attendant sympathetic increase. This sympathetic increase is not particularly surprising; it was not hypothesized, because we were basing our hypothesis partially on the results of Iwanaga and Tsukumoto (1997), but given that they used a somewhat simplistic definition of sympathetic function (i.e., LF, which is actually driven by both sympathetic and parasympathetic branches), and did not use any of the impedance cardiographic measures used in the present study, they likely would not have been able to observe an increase in sympathetic function reliably. Finally, sensation seeking modulated this effect. Although our manipulations were not identical, this contradicts the findings of Nater, Krebs, and Ehlert (2005), in that a construct related to sensation seeking was shown to influence physiological reactivity to musical stimuli.

We have strong evidence that there are differences between high and low scorers on the SSS and several of its subscales, and high and low scorers on the NCS. Looking at two of the strongest examples, DIS is predictive of how LVET changes differently from baseline to trial in the simple and complex conditions. Specifically, LVET remains unchanged in both conditions for low-DIS participants, but increases in the simple and decreases in the complex condition for high-DIS participants. Likewise, NC predicts differences in PEP reactivity, such that low-NC participants show a decrease in PEP to simple music but no change to complex, while high-NC participants show an increase to simple music but again no change to complex. However, given the absence of significant predictive power of any of the psychometric variables on any of the HRV measures, the interpretation of these results is made virtually impossible.

Although it is true that complexity was confounded with excitingness, happiness, and even liking in the present experiment, it is hard to imagine an operationalization of complexity that would be completely independent of each of these features. Given that there was a high degree of correlation between number of notes per unit time and complexity ratings, the interpretation of the manipulation as being relevant (at least) to complexity is a valid one. It therefore appears that complexity is a valid surface trait of musical stimuli, at least in the operationalization of complexity employed in this experiment. Because care was taken not to confound dissonance, tempo, or intensity with complexity, we can be sure that our results do not reflect the influence of any of these traits. Work by Haas, Distenfeld, and Axen (1986) implicated rhythm in respiratory pattern changes, but they used a tapping paradigm, and further, their definition of rhythm is more appropriately called tempo. In music theoretical terms, rhythm is the distribution of beats within a particular metrical framework, while tempo is the pace of that metrical framework, but Haas et al. do not distinguish carefully between these two. No other work has specifically studied the influence of rhythmic variations on physiology, but we assume that any variations that do not alter the complexity (see Shmulevich & Povel, 2000, for a discussion of rhythm and complexity) will have little effect on physiology.

Our results further demonstrate that complexity does have an effect on physiology, as suggested by Berlyne's (1970) hypothesis positing the existence of an inverted-U-shaped curve relationship between complexity and liking, itself related to the Yerkes-Dodson hypothesis (Yerkes & Dodson, 1908; cited in North & Hargreaves, 1995). Berlyne's hypothesis proposed "that aesthetic judgments reflect our attempts to optimize our psychobiological arousal level, which is directly related to hedonic tone, or liking" (North & Hargreaves, 1995), so any account of his hypothesis that did not consider some measure of psychobiological arousal would necessarily be incomplete. The finding in this experiment that complexity is related to psychophysiological responses, namely the withdrawal of vagal control in the autonomic nervous system, provides a mechanism to explain the results of many of the studies that have investigated the relationship between complexity and liking, as well as other relationships, e.g. between complexity and waiting time (North & Hargreaves, 1999).

The influence of sensation seeking tendencies on psychophysiological responses does not seem to provide support for the hypothesis (not explicitly tested here) that high sensation seekers enjoy more complex stimuli because their curvilinear function relating complexity and liking is shifted to the right as compared with low sensation seekers. One potential direction of the second hypothesis in our study was that a piece that physiologically overstimulates a low sensation seeker might optimally stimulate a high sensation seeker, leading to greater liking for high sensation-seekers than low sensation-seekers. However, our results seem to suggest the alternative, namely that high sensation-seekers engage the music more and therefore experience more arousal, which to them is likeable, while low sensation-seekers do not seem to engage the music as much.

Given the prevalence of music in all domains, including work and study environments, this has implications for the use of music as a means to increase productivity. For instance, Furnham and Bradley (1997) reported that introverts and extraverts were differentially distracted by background music on a cognitive test; though it was not examined here, it is probable that introversion and extraversion, being related with sensation seeking (Zuckerman, 1994) are also related to sensation seeking to some degree. In line with this prediction, Furnham and Allass (1999) found that there was a significant interaction between musical complexity and extroversion, such that increasing-complexity musical distraction resulted in an increase in extroverts', and a decrease in introverts', cognitive task performance. Kiger (1989) examined the influence of complexity on a cognitive task more explicitly, and found a significant effect for complexity, but did not examine any constructs like sensation seeking. Clearly, the results of the present experiment are in line with the findings of both of the experiments just discussed, and suggest an explanation for the patterns of results found between the two.

Avenues for future research exist in two directions. First, following Berlyne, the construct of sensation seeking should be examined, using a broader array of stimuli and more thorough statistical analyses than could be done here. Second, the influence of complexity on task performance and in musical preference studies should be examined using the psychophysiological methods laid out in this experiment, in order to more precisely determine the physiological mechanisms underlying the patterns of results obtained.

	Key	Avg. pitch wtd/unwtd	Avg. dist. b/tw notes	Range	Tempo	Notes/mm	# mm	Length
Allemande								
Complex	A	16.37/16.34	3.52	F#3-A5	86	15	32	1:32
Simple	A	16.49/16.55	4.81	F#3-A5	86	8	32	1:32
Bravade								
Complex	d	23.40/23.25	2.43	D4-A5	83	11.25	32	1:36
Simple	d	23.02/23.05	2.56	D4-A5	83	4.19	32	1:36
Engels Nachtegaeltje								
Complex	C	21.37/21.54	2.61	C4-A5	94	12.62	42	1:48
Simple	C	21.63/21.81	4.30	C4-A5	94	5.38	42	1:48
Kits Allemande								
Complex	b	19.90/19.82	2.66	B3-A5	92	11.64	44	1:58
Simple	b	21.35/21.29	3.39	B3-A5	92	4.30	44	1:58
Moderato								
Complex	Bb	15.68/15.95	3.21	F3-Bb5	96	9.22	41	1:43
Simple	Bb	16.14/15.95	4.19	F3-Bb5	96	4.63	41	1:43

Table 1A. Titles and descriptions of each piece used. Abbreviations: avg. = average; wtd = weighted; unwtd = unweighted; b/tw = between; mm = measure.

<u>Pitch :: number equivalents</u>

F3 = 1	G#4 = 16
F#3 = 2	A4 = 17
G3 = 3	A#4 = 18
G#3 = 4	B4 = 19
A3 = 5	C5 = 20
A#3 = 6	C#5 = 21
B3 = 7	D5 = 22
C4 = 8	D#5 = 23
C#4 = 9	E5 = 24
D4 = 10	F5 = 25
D#4 = 11	F#5 = 26
E4 = 12	G5 = 27
F4 = 13	G#5 = 28
F#4 = 14	A5 = 29
G4 = 15	A#5 = 30

Piece	Composer	Full title
Allemande	J.S. Bach (ed. Marsteller)	BWV 1007, 2nd mvt.
Bravade	Jakob van Eyck	Der Fluiten Lusthof - Bravade; Variatie 2
Engels Nachtegaeltje	J. van Eyck	DFL - Engels nachtegaeltje; Modo 3
Kits Allemande	J. van Eyck	DFL - Kits alamande; Variatie 1
Moderato	O. Blume (ed. R. Fink)	36 Studies for Trombone, # 18

Table 1B. Additional information for each piece.

	ss	tas	es	dis	bs	nc	pc	pcs	pcd	mps	stai	staidif
N	30	30	30	30	30	30	30	30	30	30	30	30
Mean	19.13	6.43	5.07	4.23	3.40	112.20	47.67	16.93	19.07	25.07	8.60	-1.33
Median	18.00	6.50	5.00	4.00	3.50	115.00	48.50	16.00	19.50	24.50	9.00	-1.00
Mode	18	10	5	3	4	119	43	14(a)	17	19	9	-2(a)
Std. Deviation	5.387	2.873	1.780	2.373	1.714	17.371	7.512	3.787	3.352	6.269	1.831	1.470
Skewness	.253	-.244	.325	-.082	.468	-.054	-.122	.353	-.297	.814	.353	.486
Std. Error of Skewness	.427	.427	.427	.427	.427	.427	.427	.427	.427	.427	.427	.427
Kurtosis	-.323	-1.197	-.225	-.504	-.382	-.515	-.311	-.752	-1.115	.184	-.688	1.310
Std. Error of Kurtosis	.833	.833	.833	.833	.833	.833	.833	.833	.833	.833	.833	.833
Minimum	10	1	2	0	1	76	32	10	13	17	6	-4
Maximum	30	10	9	9	7	150	64	24	24	40	12	3

a Multiple modes exist. The smallest value is shown

Table 2. Descriptive statistics on all subject variables.

1. gender * complexity * period Measure: CO						
gender	complexity	period	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
female	**simple**	**base**	17.525	2.115	13.170	21.880
		trial	17.250	2.093	12.939	21.561
	complex	**base**	17.188	2.165	12.730	21.646
		trial	16.934	2.117	12.573	21.294
male	**simple**	**base**	26.231	2.194	21.711	30.750
		trial	27.592	2.172	23.119	32.066
	complex	**base**	26.996	2.247	22.369	31.623
		trial	26.607	2.197	22.082	31.132

Table 3A. Statistics for CO.

2. gender * complexity * period Measure: SV						
gender	complexity	period	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
female	**simple**	**base**	257.078	32.315	190.524	323.632
		trial	257.736	31.507	192.846	322.626
	complex	**base**	252.937	31.886	187.267	318.607
		trial	254.653	31.753	189.255	320.050
male	**simple**	**base**	362.893	33.535	293.827	431.959
		trial	384.319	32.697	316.979	451.659
	complex	**base**	369.645	33.089	301.497	437.794
		trial	368.973	32.952	301.107	436.839

Table 3B. Statistics for SV.

3. gender * complexity * period Measure: HF						
gender	**complexity**	**period**	**Mean**	**Std. Error**	**95% Confidence Interval**	
					Lower Bound	**Upper Bound**
female	**simple**	**base**	6.859	.287	6.269	7.448
		trial	6.961	.274	6.397	7.525
	complex	**base**	7.033	.240	6.540	7.525
		trial	7.081	.272	6.521	7.640
male	**simple**	**base**	6.104	.267	5.555	6.652
		trial	6.044	.255	5.519	6.568
	complex	**base**	6.191	.223	5.733	6.650
		trial	5.922	.253	5.401	6.442

Table 3C. Statistics for HF.

4. gender * complexity * period Measure: RR						
gender	**complexity**	**period**	**Mean**	**Std. Error**	**95% Confidence Interval**	
					Lower Bound	**Upper Bound**
female	**simple**	**base**	15.946	.657	14.600	17.291
		trial	16.199	.651	14.866	17.532
	complex	**base**	15.942	.644	14.622	17.262
		trial	16.754	.682	15.357	18.151
male	**simple**	**base**	14.068	.614	12.810	15.327
		trial	14.348	.609	13.101	15.595
	complex	**base**	13.720	.603	12.485	14.955
		trial	14.541	.638	13.234	15.848

Table 3D. Statistics for RR.

5. gender * complexity * period Measure: LVET						
gender	**complexity**	**period**	**Mean**	**Std. Error**	**95% Confidence Interval**	
					Lower Bound	**Upper Bound**
female	**simple**	**base**	275.125	10.226	254.178	296.072
		trial	269.174	11.347	245.930	292.417
	complex	**base**	276.781	10.982	254.285	299.278
		trial	268.232	12.348	242.938	293.526
male	**simple**	**base**	256.276	9.565	236.682	275.869
		trial	267.299	10.614	245.556	289.041
	complex	**base**	257.187	10.273	236.143	278.231
		trial	256.086	11.551	232.425	279.746

Table 3E. Statistics for LVET.

6. gender * complexity * period Measure: LF						
gender	**complexity**	**period**	**Mean**	**Std. Error**	**95% Confidence Interval**	
					Lower Bound	**Upper Bound**
female	**simple**	**base**	6.950	.198	6.543	7.357
		trial	6.608	.194	6.210	7.007
	complex	**base**	7.075	.188	6.690	7.461
		trial	6.810	.198	6.405	7.216
male	**simple**	**base**	6.936	.192	6.543	7.330
		trial	6.676	.187	6.291	7.060
	complex	**base**	7.042	.181	6.670	7.414
		trial	6.697	.191	6.305	7.089

Table 3F. Statistics for LF.

7. gender * complexity * period Measure: LH						
gender	complexity	period	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
female	simple	base	1.465	.286	.875	2.056
		trial	1.013	.197	.607	1.418
	complex	base	1.547	.287	.955	2.139
		trial	1.102	.236	.615	1.589
male	simple	base	2.421	.309	1.784	3.059
		trial	1.905	.212	1.467	2.344
	complex	base	2.475	.310	1.836	3.115
		trial	1.999	.255	1.473	2.525

Table 3G. Statistics for LH.

8. gender * complexity * period Measure: IBI						
gender	complexity	period	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
female	simple	base	880.314	29.356	820.182	940.446
		trial	897.117	29.780	836.116	958.119
	complex	base	889.913	27.197	834.203	945.623
		trial	900.877	27.562	844.418	957.336
male	simple	base	838.299	27.460	782.051	894.548
		trial	842.303	27.857	785.242	899.365
	complex	base	832.712	25.440	780.600	884.824
		trial	840.994	25.782	788.181	893.806

Table 3H. Statistics for IBI.

1. sssplit * complexity * period Measure: LVET						
sssplit	**complexity**	**period**	**Mean**	**Std. Error**	**95% Confidence Interval**	
			Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
low ss (N=10)	**simple**	**base**	271.137	12.954	244.023	298.250
		trial	268.856	12.629	242.423	295.289
	complex	**base**	265.663	13.592	237.215	294.112
		trial	266.448	14.800	235.470	297.426
high ss (N=11)	**simple**	**base**	272.375	12.351	246.524	298.227
		trial	274.452	12.041	249.249	299.655
	complex	**base**	276.375	12.960	249.251	303.500
		trial	266.509	14.112	236.973	296.045

Table 4A. Statistics for LVETxSS.

2. sssplit * complexity * period Measure: SV						
sssplit	**complexity**	**period**	**Mean**	**Std. Error**	**95% Confidence Interval**	
			Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
low ss (N=10)	**simple**	**base**	290.013	43.956	197.274	382.752
		trial	294.900	42.811	204.578	385.223
	complex	**base**	284.990	42.371	195.594	374.386
		trial	294.923	43.060	204.076	385.771
high ss (N=9)	**simple**	**base**	354.534	46.334	256.778	452.289
		trial	359.233	45.127	264.024	454.442
	complex	**base**	358.646	44.663	264.415	452.877
		trial	350.747	45.389	254.985	446.509

Table 4B. Statistics for SVxSS.

3. sssplit * complexity * period Measure: PEP						
sssplit	**complexity**	**period**	**Mean**	**Std. Error**	**95% Confidence Interval**	
			Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
low ss (N=10)	**simple**	**base**	103.457	2.502	98.179	108.735
		trial	103.709	2.407	98.630	108.788
	complex	**base**	103.404	1.962	99.264	107.544
		trial	104.279	1.945	100.176	108.382
high ss (N=9)	**simple**	**base**	99.461	2.133	94.960	103.962
		trial	99.787	2.053	95.456	104.118
	complex	**base**	99.061	1.674	95.530	102.592
		trial	98.709	1.658	95.210	102.208

Table 4C. Statistics for PEPxSS.

4. essplit * complexity * period Measure: LVET						
essplit	**complexity**	**period**	**Mean**	**Std. Error**	**95% Confidence Interval**	
			Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
low es (N=11)	**simple**	**base**	287.870	10.753	265.439	310.301
		trial	292.394	11.223	268.982	315.806
	complex	**base**	286.048	11.739	261.560	310.536
		trial	290.232	12.406	264.354	316.110
high es (N=11)	**simple**	**base**	251.951	10.753	229.520	274.382
		trial	260.487	11.223	237.075	283.898
	complex	**base**	262.713	11.739	238.226	287.201
		trial	252.120	12.406	226.242	277.998

Table 4D. Statistics for LVETxES.

5. essplit * complexity * period Measure: SV						
essplit	**complexity**	**period**	**Mean**	**Std. Error**	**95% Confidence Interval**	
			Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
low es (N=10)	**simple**	**base**	310.939	43.514	219.863	402.016
		trial	320.003	43.219	229.544	410.462
	complex	**base**	311.408	43.105	221.188	401.628
		trial	324.516	42.526	235.508	413.523
high es (N=11)	**simple**	**base**	350.070	41.489	263.232	436.908
		trial	366.051	41.208	279.802	452.300
	complex	**base**	355.209	41.099	269.188	441.231
		trial	346.104	40.547	261.239	430.969

Table 4E. Statistics for SVxES.

6. essplit * complexity * period Measure: CO						
essplit	**complexity**	**period**	**Mean**	**Std. Error**	**95% Confidence Interval**	
			Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
low es (N=10)	**simple**	**base**	20.954	2.885	14.915	26.993
		trial	21.467	2.946	15.301	27.633
	complex	**base**	21.267	2.991	15.006	27.528
		trial	21.865	2.953	15.684	28.047
high es (N=11)	**simple**	**base**	25.609	2.751	19.852	31.367
		trial	26.504	2.809	20.624	32.383
	complex	**base**	26.010	2.852	20.041	31.980
		trial	24.943	2.816	19.049	30.836

Table 4F. Statistics for COxES.

7. dissplit * complexity * period Measure: LVET						
dissplit	complexity	period	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound	Lower Bound	Upper Bound
low dis (N=12)	simple	base	260.561	11.921	235.956	285.165
		trial	260.253	12.728	233.984	286.522
	complex	base	256.479	12.648	230.374	282.583
		trial	259.175	14.202	229.862	288.487
high dis (N=14)	simple	base	266.214	11.037	243.434	288.993
		trial	274.021	11.784	249.701	298.341
	complex	base	272.985	11.710	248.817	297.153
		trial	261.771	13.149	234.634	288.909

Table 4G. Statistics for LVETxDIS.

8. dissplit * complexity * period Measure: SV						
dissplit	complexity	period	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound	Lower Bound	Upper Bound
low dis (N=11)	simple	base	264.602	39.002	183.493	345.711
		trial	267.898	38.427	187.984	347.812
	complex	base	257.034	38.241	177.507	336.561
		trial	267.152	38.537	187.011	347.293
high dis (N=12)	simple	base	341.986	37.342	264.329	419.642
		trial	362.582	36.791	286.070	439.094
	complex	base	351.058	36.613	274.916	427.199
		trial	345.522	36.896	268.792	422.251

Table 4H. Statistics for SVxDIS.

9. dissplit * complexity * period Measure: CO						
dissplit	**complexity**	**period**	**Mean**	**Std. Error**	**95% Confidence Interval**	
			Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
low dis (N=11)	**simple**	**base**	18.936	2.714	13.293	24.579
		trial	18.985	2.783	13.199	24.772
	complex	**base**	18.445	2.802	12.618	24.273
		trial	18.691	2.765	12.942	24.440
high dis (N=12)	**simple**	**base**	23.939	2.598	18.536	29.342
		trial	25.183	2.664	19.643	30.724
	complex	**base**	24.819	2.683	19.240	30.398
		trial	24.145	2.647	18.640	29.649

Table 4I. Statistics for COxDIS.

10. dissplit * complexity * period Measure: PEP						
dissplit	**complexity**	**period**	**Mean**	**Std. Error**	**95% Confidence Interval**	
			Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
low dis (N=9)	**simple**	**base**	103.724	2.285	98.972	108.475
		trial	103.366	2.188	98.816	107.917
	complex	**base**	103.920	1.894	99.981	107.859
		trial	103.927	1.931	99.911	107.943
high dis (N=14)	**simple**	**base**	98.758	1.832	94.948	102.567
		trial	99.167	1.754	95.518	102.815
	complex	**base**	98.144	1.519	94.986	101.302
		trial	98.143	1.548	94.923	101.363

Table 4J. Statistics for PEPxDIS.

11. bssplit * complexity * period Measure: IBI						
bssplit	**complexity**	**period**	**Mean**	**Std. Error**	**95% Confidence Interval**	
			Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
low bs (N=11)	**simple**	**base**	860.656	32.121	794.361	926.950
		trial	868.709	32.266	802.115	935.303
	complex	**base**	857.783	28.645	798.662	916.904
		trial	870.652	28.649	811.524	929.780
high bs (N=15)	**simple**	**base**	836.137	27.507	779.365	892.908
		trial	847.811	27.631	790.783	904.839
	complex	**base**	838.019	24.530	787.391	888.647
		trial	846.859	24.533	796.225	897.493

Table 4K. Statistics for IBIxBS.

12. ncsplit * complexity * period Measure: PEP						
ncsplit	**complexity**	**period**	**Mean**	**Std. Error**	**95% Confidence Interval**	
			Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**
low nc (N=10)	**simple**	**base**	103.091	2.498	97.820	108.362
		trial	102.016	2.423	96.903	107.129
	complex	**base**	101.702	2.083	97.308	106.097
		trial	102.165	2.145	97.640	106.690
high nc (N=9)	**simple**	**base**	99.713	2.634	94.157	105.269
		trial	100.516	2.555	95.127	105.906
	complex	**base**	99.417	2.196	94.785	104.049
		trial	99.421	2.261	94.651	104.191

Table 4L. Statistics for PEPxNC.

Variable pair	Multivariate test (covariate)	Multivariate test (high-low)	Correlation
SS-LVET	$F(1,27) = 6.98$ **$p = .014$**	$F(1,17) = 4.10$ $p = .059$	$r = -.495$ **$p = .005$**
SS-SV	$F(1,24) = 2.66$ $p = .116$	$F(1,15) = 2.05$ $p = .172$	$r = -.398$ **$p = .040$**
SS-PEP	$F(1,24) = 4.13$ $p = .053$	$F(1,15) = 1.39$ $p = .257$	$r = -.284$ $p = .152$
ES-LVET	$F(1,27) = 8.19$ **$p = .008$**	$F(1,18) = 4.36$ $p = .051$	$r = -.484$ **$p = .007$**
ES-SV	$F(1,24) = 10.71$ **$p = .003$**	$F(1,17) = 4.68$ $p = .045$	$r = -.545$ **$p = .003$**
ES-CO	$F(1,24) = 13.01$ **$p = .001$**	$F(1,17) = 4.86$ $p = .042$	$r = -.570$ **$p = .002$**
DIS-LVET	$F(1,27) = 4.61$ $p = .041$	$F(1,22) = 7.92$ **$p = .010$**	$r = -.414$ **$p = .023$**
DIS-SV	$F(1,24) = 2.31$ $p = .141$	$F(1,19) = 7.48$ **$p = .013$**	$r = -.338$ $p = .085$
DIS-CO	$F(1,24) = 1.12$ $p = .301$	$F(1,19) = 6.18$ **$p = .022$**	$r = -.265$ $p = .182$
DIS-PEP	$F(1,24) = 1.92$ $p = .179$	$F(1,19) = 0.88$ $p = .359$	$r = -.170$ $p = .398$
BS-IBI	$F(1,27) = 2.82$ $p = .105$	$F(1,22) = 0.85$ $p = .397$	$r = -.300$ $p = .107$
NC-PEP	$F(1,24) = 7.91$ **$p = .010$**	$F(1,15) = 10.84$ **$p = .005$**	$r = -.594$ **$p = .001$**
STAI-LVET	$F(1,27) = 19.83$ **$p = .000$**	$F(1,22) = 3.61$ $p = .071$	$r = -.671$ **$p = .000$**
STAI-SV	$F(1,24) = 8.26$ **$p = .008$**	$F(1,20) = 1.65$ $p = .214$	$r = -.610$ **$p = .001$**
STAI-CO	$F(1,24) = 10.64$ **$p = .003$**	$F(1,20) = 2.07$ $p = .166$	$r = -.661$ **$p = .000$**
STAIDIF-PEP	$F(1,24) = 2.09$ $p = .161$	$F(1,10) = 4.65$ $p = .056$	$r = .080$ $p = .692$

Table 5A. All models used for multivariate tests include gender.

Variable pair	Multivariate test (covariate)	Multivariate test (high-low)	Significant t-tests	*p* values
SS-LVET	$F(1,28) = 9.10$ ***p* = .005**	$F(1,19) = 4.76$ $p = .042$	N/A	N/A
SS-SV	$F(1,25) = 4.70$ $p = .040$	$F(1,17) = 3.17$ $p = .093$	ss*c*p1	.008
SS-PEP	$F(1,25) = 3.69$ $p = .066$	$F(1,17) = 1.94$ $p = .182$	lss*c ss*s*p1 ss*s*p2 ss*c*p1 ss*c*p2	.005 .006 .004 .001 .000
ES-LVET	$F(1,28) = 8.58$ ***p* = .007**	$F(1,20) = 4.77$ $p = .041$	ss*s*p1 ss*s*p2 ss*c*p1 ss*c*p2	.000 .000 .000 .000
ES-SV	$F(1,25) = 10.56$ ***p* = .003**	$F(1,19) = 5.37$ $p = .032$	N/A	N/A
ES-CO	$F(1,25) = 12.05$ ***p* = .002**	$F(1,19) = 5.86$ ***p* = .026**	N/A	N/A
DIS-LVET	$F(1,28) = 5.78$ ***p* = .023**	$F(1,24) = 10.13$ ***p* = .004**	N/A	N/A
DIS-SV	$F(1,25) = 3.23$ $p = .085$	$F(1,21) = 9.20$ ***p* = .006**	ss*s*p1 ss*s*p2 ss*c*p1 ss*c*p2	.005 .001 .001 .006
DIS-CO	$F(1,25) = 1.89$ $p = .182$	$F(1,21) = 7.55$ ***p* = .012**	ss*s*p1 ss*s*p2 ss*c*p1 ss*c*p2	.008 .002 .001 .005
DIS-PEP	$F(1,25) = 1.88$ $p = .182$	$F(1,21) = 0.77$ $p = .390$	ss*s*p1 ss*s*p2 ss*c*p1 ss*c*p2	.001 .001 .000 .000
BS-IBI	$F(1,28) = 2.77$ $p = .107$	$F(1,24) = 1.28$ $p = .268$	hss*s lss*c	.012 .004
NC-PEP	$F(1,25) = 8.20$ ***p* = .008**	$F(1,17) = 9.21$ ***p* = .007**	N/A	N/A
STAI-LVET	$F(1,28) = 22.93$ ***p* = .000**	$F(1,24) = 5.90$ ***p* = .023**	ss*s*p1	.002
STAI-SV	$F(1,25) = 14.85$ ***p* = .001**	$F(1,22) = 3.96$ $p = .059$	hss*s	.011
STAI-CO	$F(1,25) = 19.38$ ***p* = .000**	$F(1,22) = 4.61$ $p = .043$	N/A	N/A
STAIDIF-PEP	$F(1,25) = 1.67$ $p = .208$	$F(1,12) = 9.04$ ***p* = .011**	N/A	N/A

Table 5B. All models used for multivariate tests collapse across gender.

dv	C	E	R*	H	S	L**
hf	0.09	0.04	0.20	0.06	0.07	0.49
lf	0.81	0.80	0.53	0.82	0.81	0.85
lh	0.00	0.00	0.27	0.09	0.08	0.24
rr	0.60	0.63	0.37	0.54	0.56	0.56
lv	0.63	0.59	0.22	0.40	0.35	0.58
sv	0.21	0.11	0.31	0.08	0.16	0.19
sv2	0.93	0.92	0.61	0.90	0.91	0.77
co	0.15	0.05	0.47	0.05	0.01	0.20
co2	0.90	0.91	0.41	0.76	0.72	0.84
pep	0.21	0.11	0.06	0.15	0.21	0.42
ibi	0.00	0.12	0.06	0.02	0.04	0.17

Table 6. Correlations between mean variable ratings. All values given are *r* values between averages of that dependent variable (dv) for each song, and averages of the MEQ question indicated for each song.

*Note: these are all artificially high because of the narrow range of **R**.
Same as **R.
C = complexity
E = excitingness
R = relaxingness
H = happiness
S = sadness
L = liking

sv2 and co2 are corrected version of sv and co with one significant outlier removed.

Figure 1A. Unornamented version of mm. 1-4 of *Allemande*.

Figure 1B. Ornamented version of mm. 1-4 of *Allemande*.

The principle approach taken to creating multiple versions of each piece was to use pre-existing theme and variation sets; this was the case for *Bravade*, *Engels Nachtegaeltje*, and *Kits Allemande*. However, for two of the pieces (*Allemande*, shown above, and *Moderato*), a "reverse-engineering" approach was taken – that is to say, I started with a complex piece and derived a simple version by retaining the harmonic structure while eliminating many passing and neighbor tones.

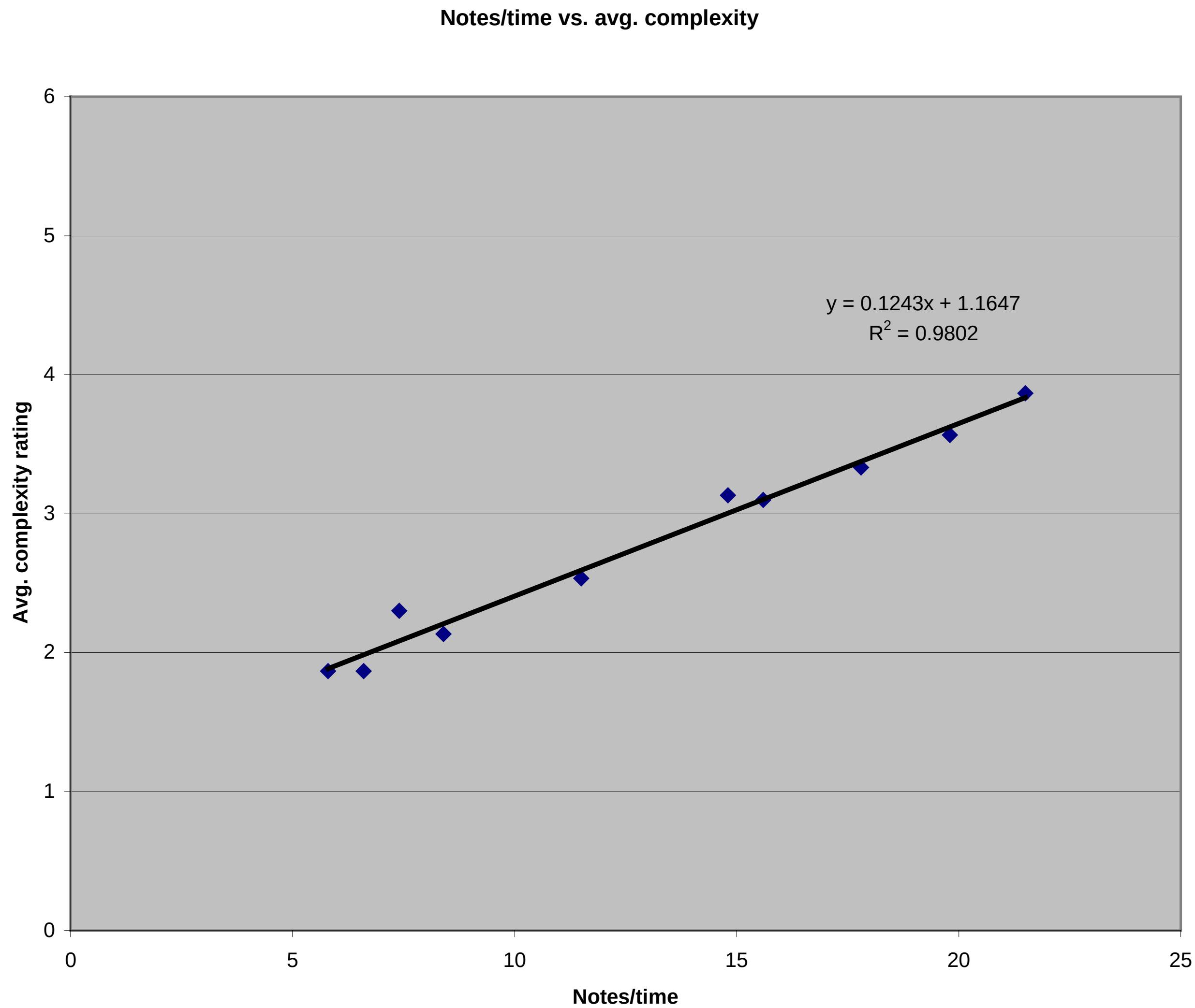

Figure 2A. Average complexity ratings plotted against the number of notes per unit time for each song.

$y = 0.1243x + 1.1647$
$R^2 = 0.9802$

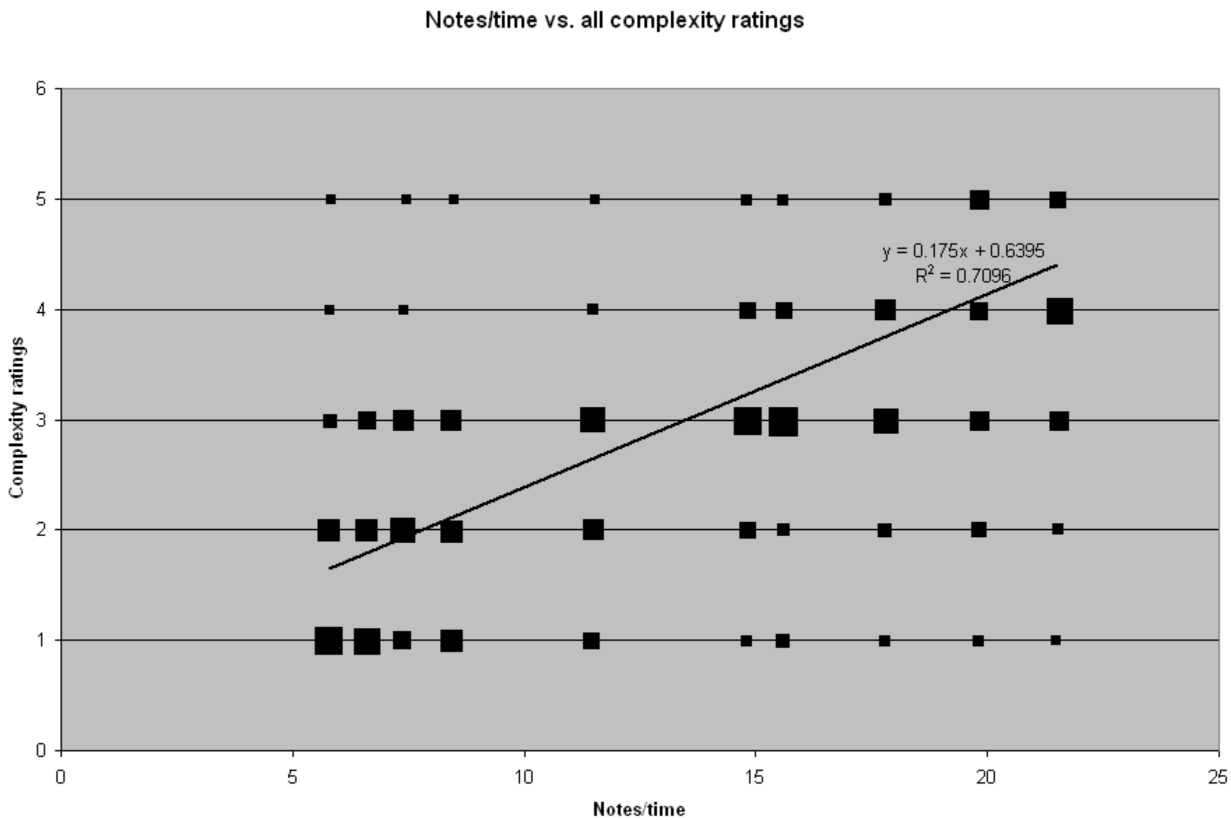

Figure 2B. Raw complexity ratings (size of square is indicative of number of participants who chose that complexity rating) plotted against the number of notes per unit time for each song.

$y = 0.175x + 0.6395$
$R^2 = 0.7096$

<u>Key for Figures 3-10</u>
complexity 1 = simple
complexity 2 = complex
period 1 = baseline preceding song
period 2 = song

Note: All ordinates for female and male graphs span the same distance.

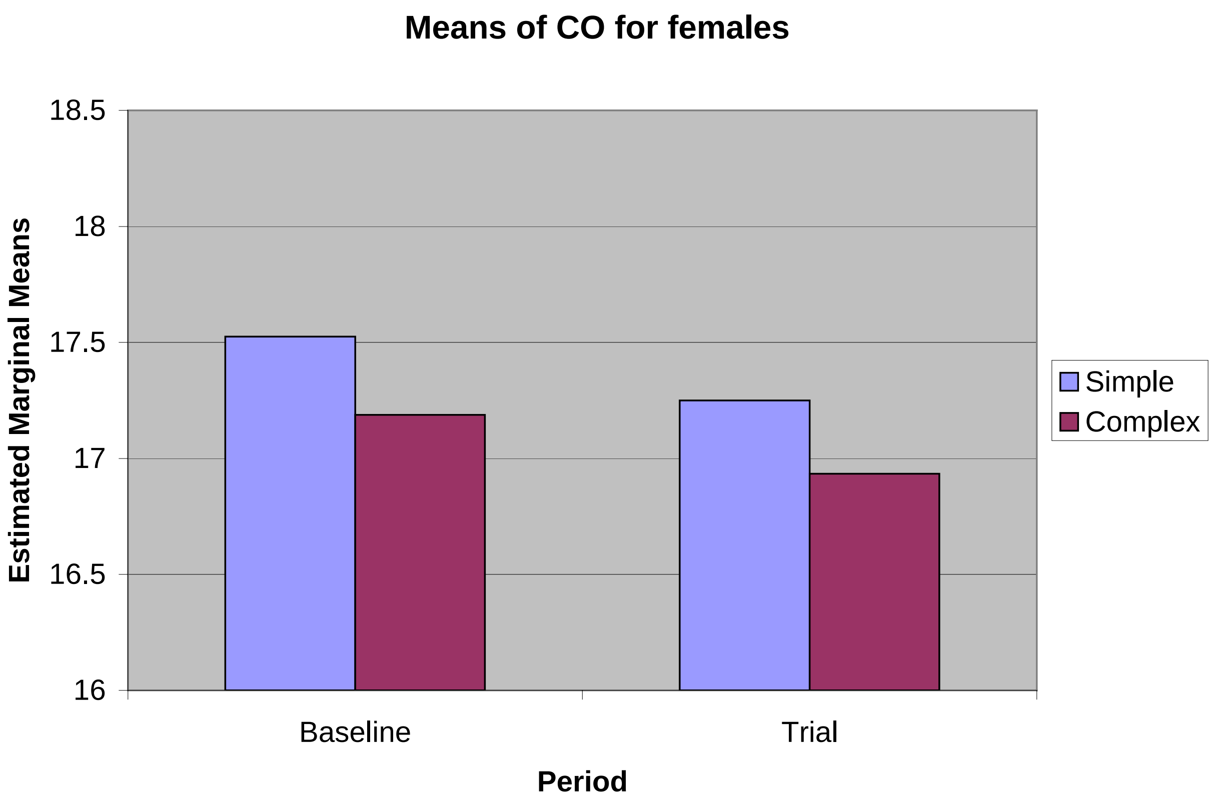

Figure 3A. Means for CO for females, collapsed across songs.

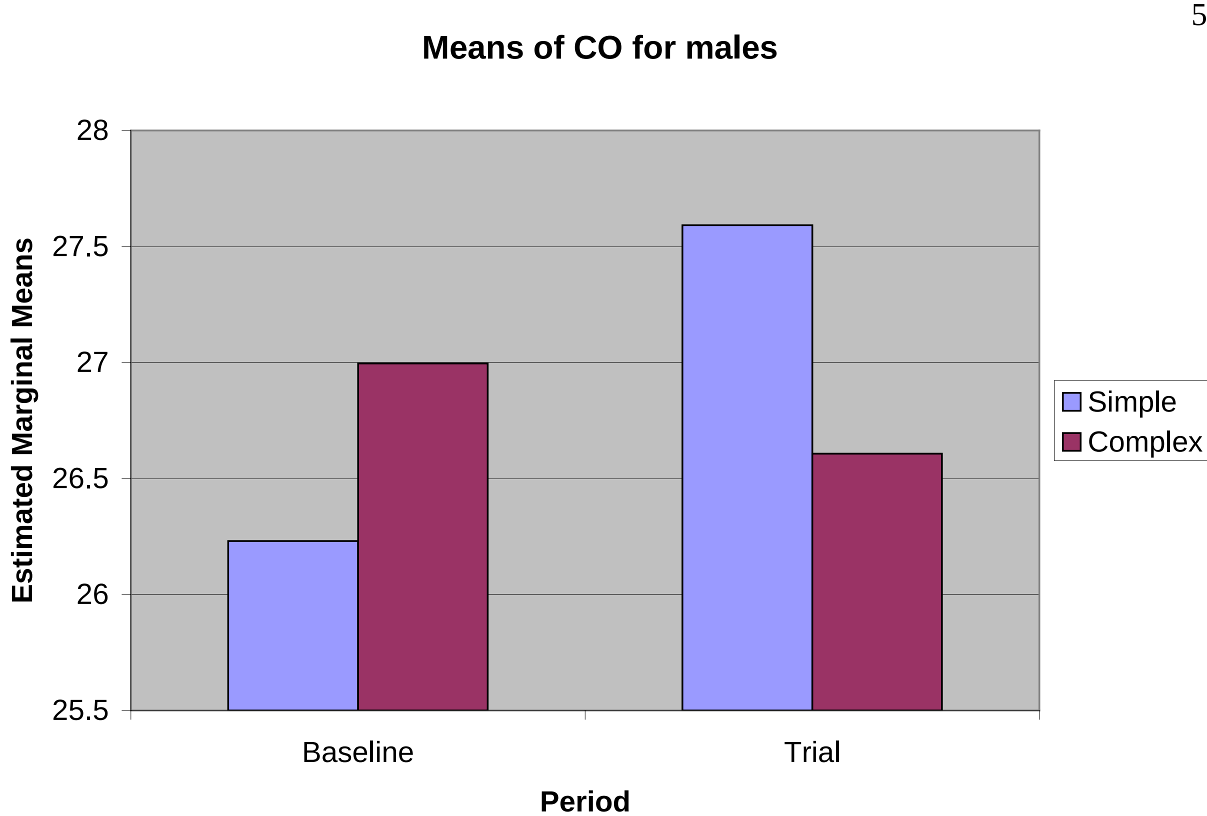

Figure 3B. Means for CO for males, collapsed across songs.

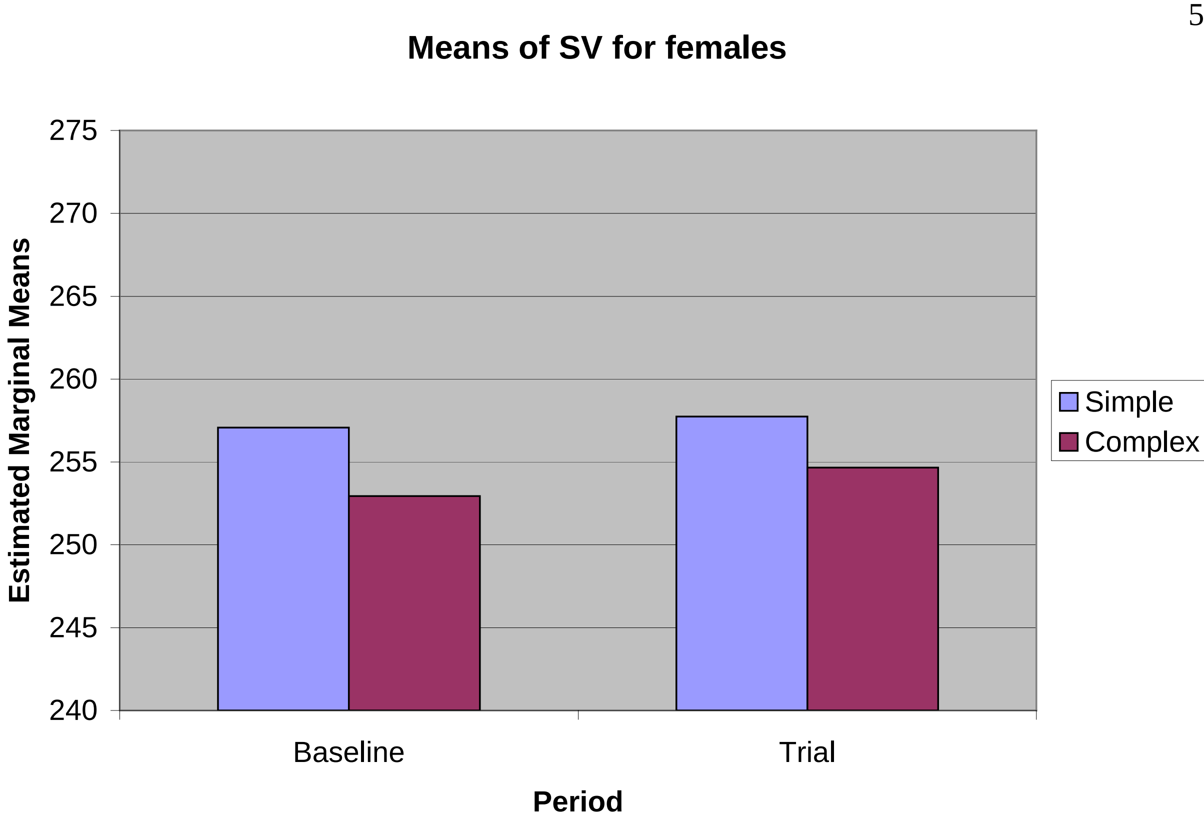

Figure 4A. Means for SV for females, collapsed across songs.

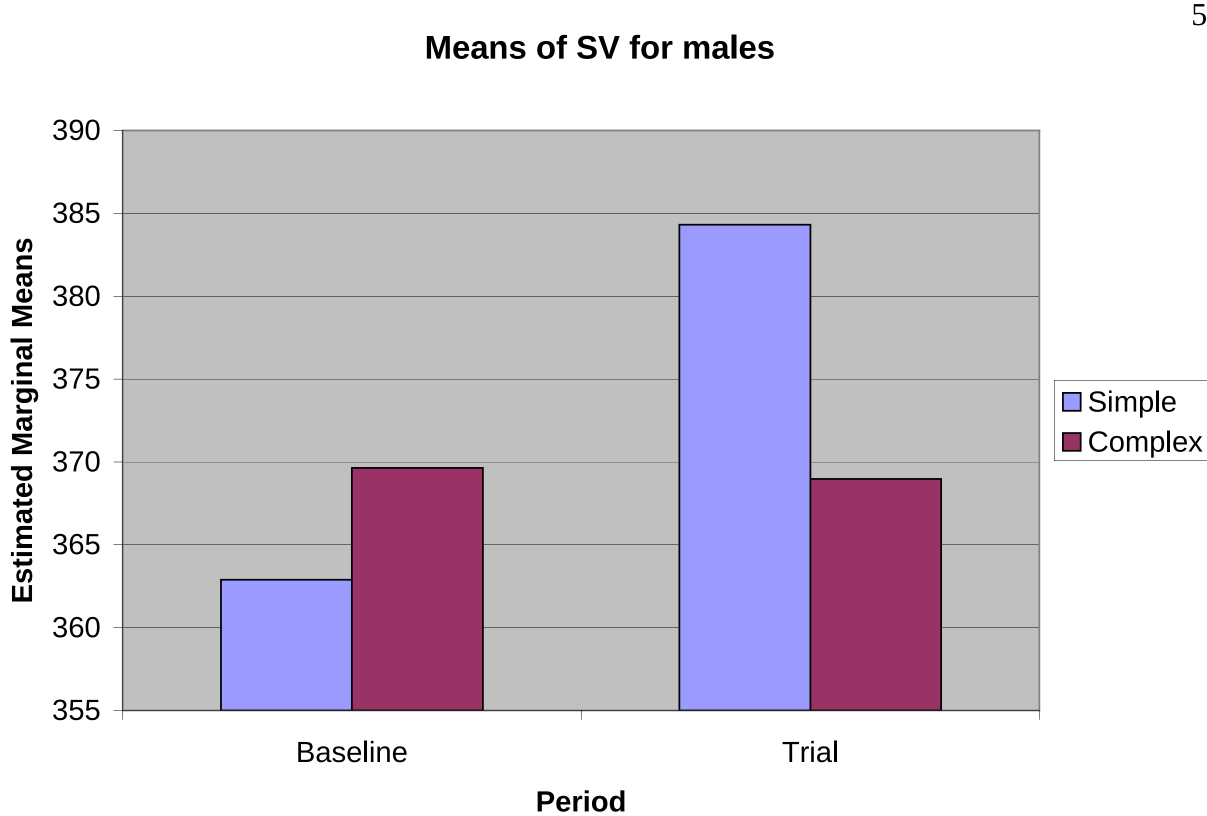

Figure 4B. Means for SV for males, collapsed across songs.

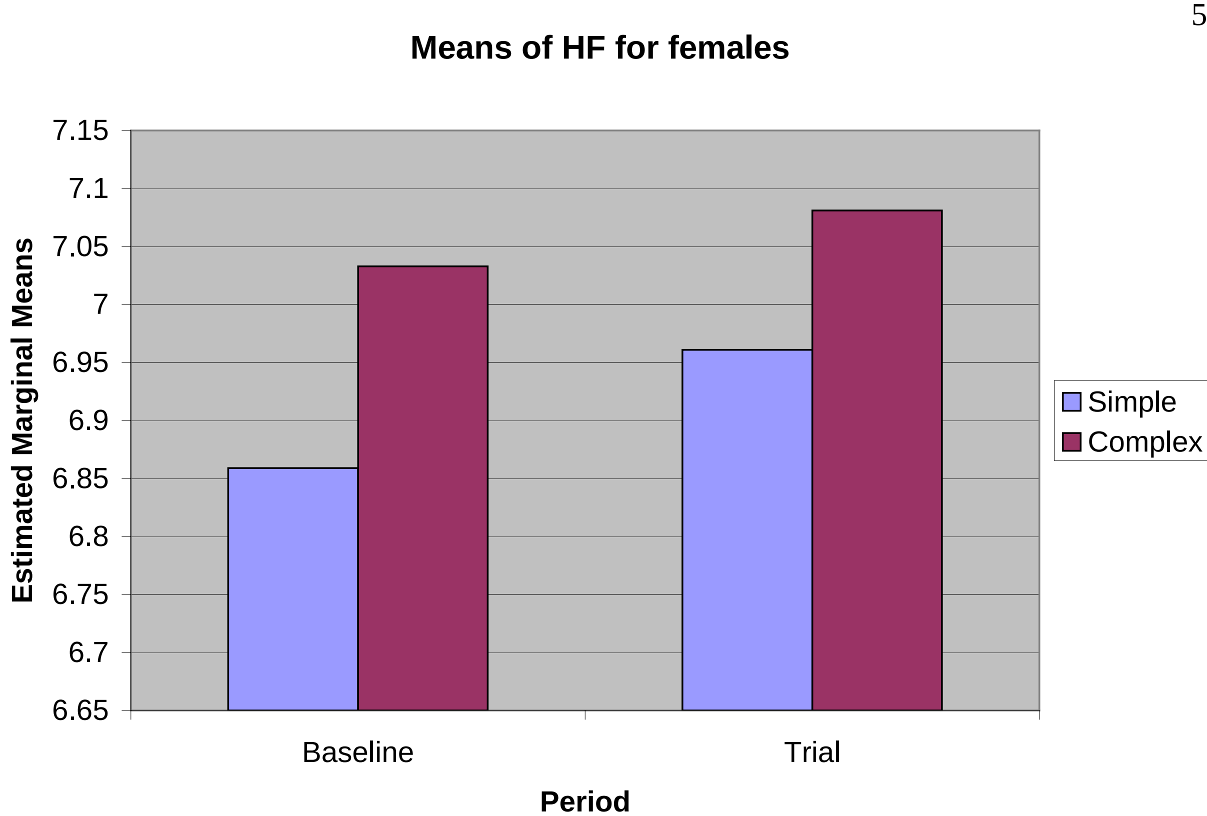

Figure 5A. Means for HF for females, collapsed across songs.

Figure 5B. Means for HF for males, collapsed across songs.

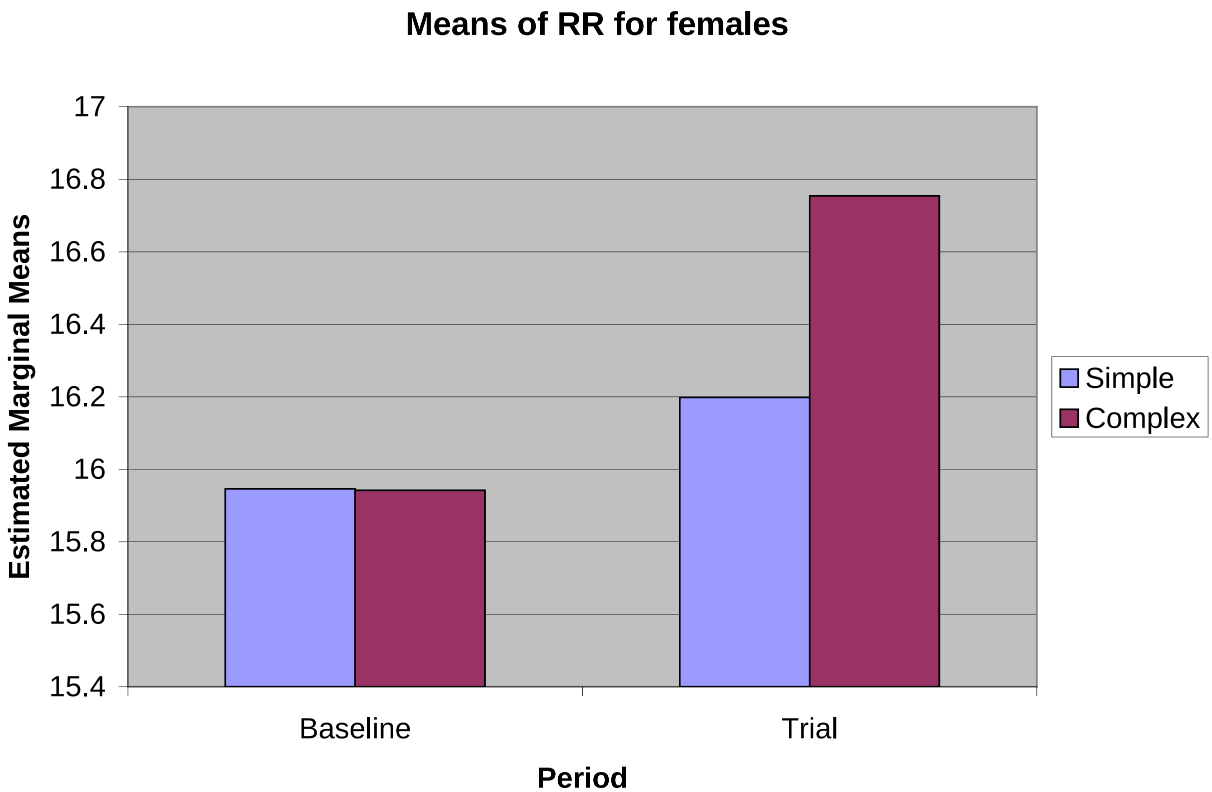

Figure 6A. Means for RR for females, collapsed across songs.

Figure 6B. Means for RR for males, collapsed across songs.

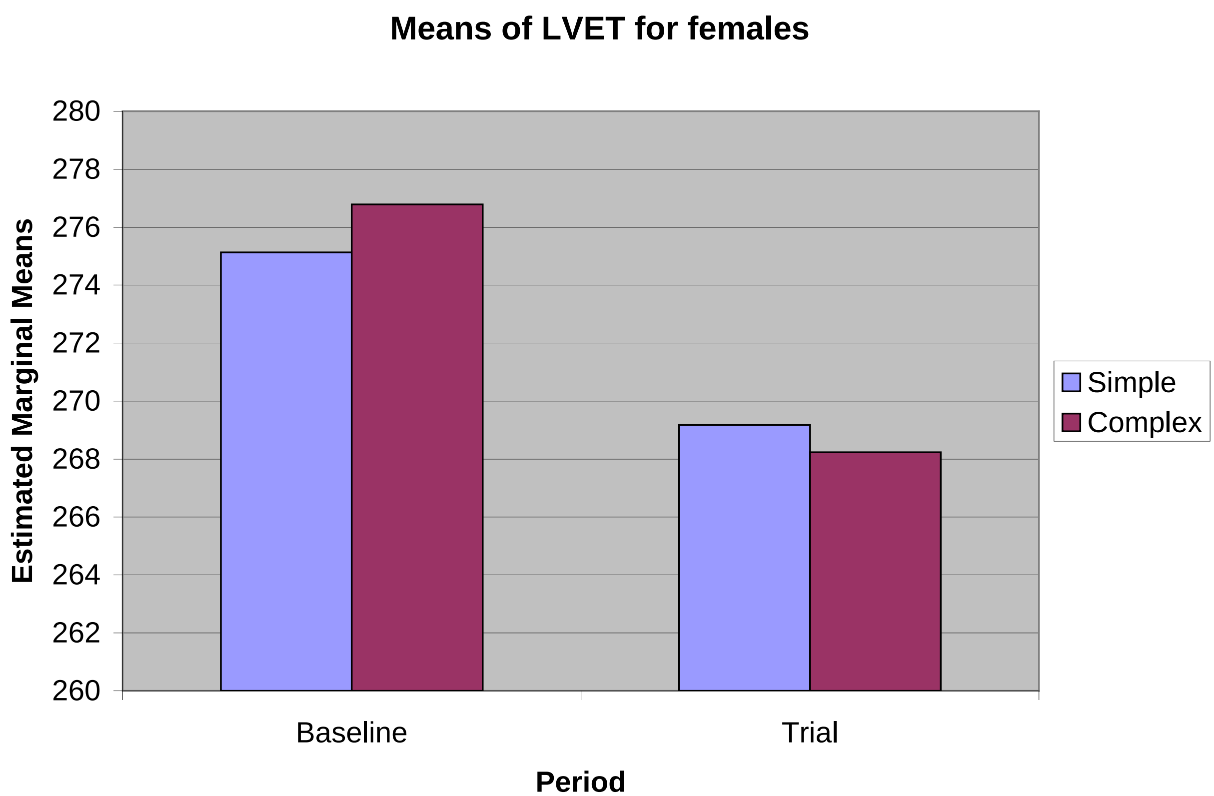

Figure 7A. Means for LVET for females, collapsed across songs.

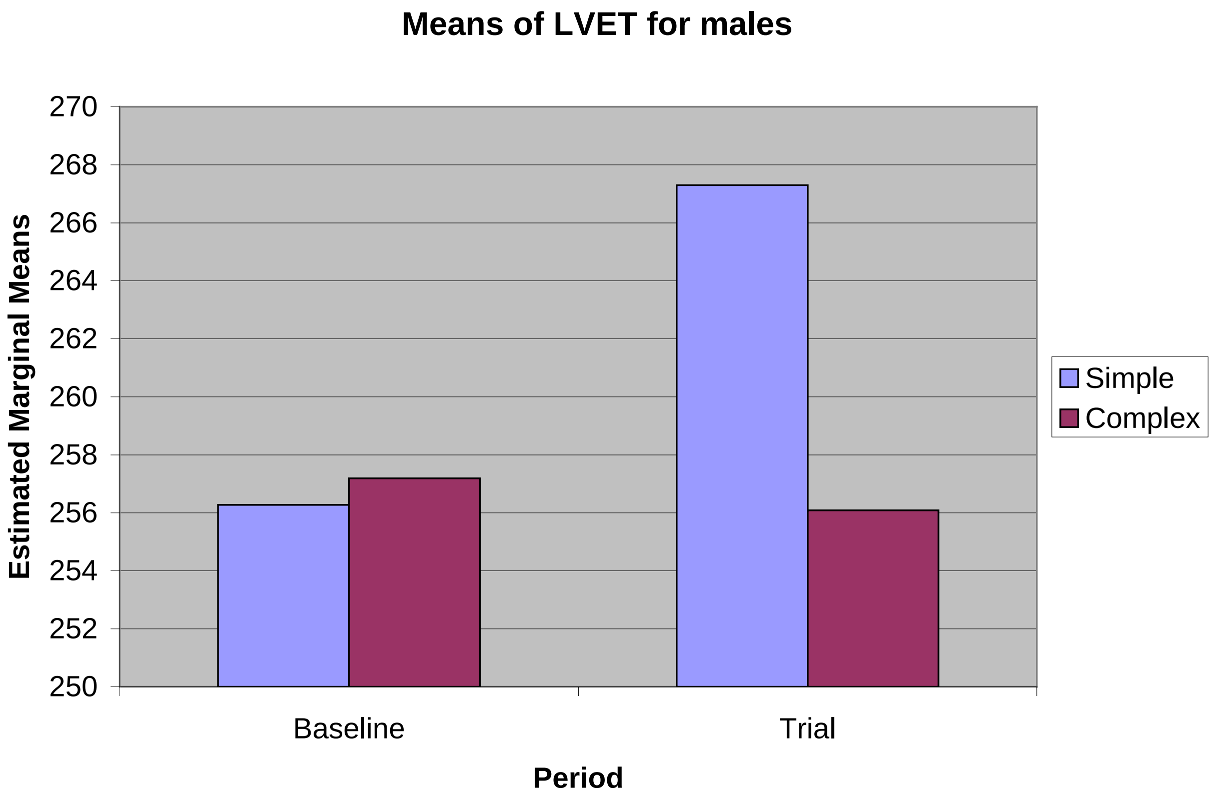

Figure 7B. Means for LVET for males, collapsed across songs.

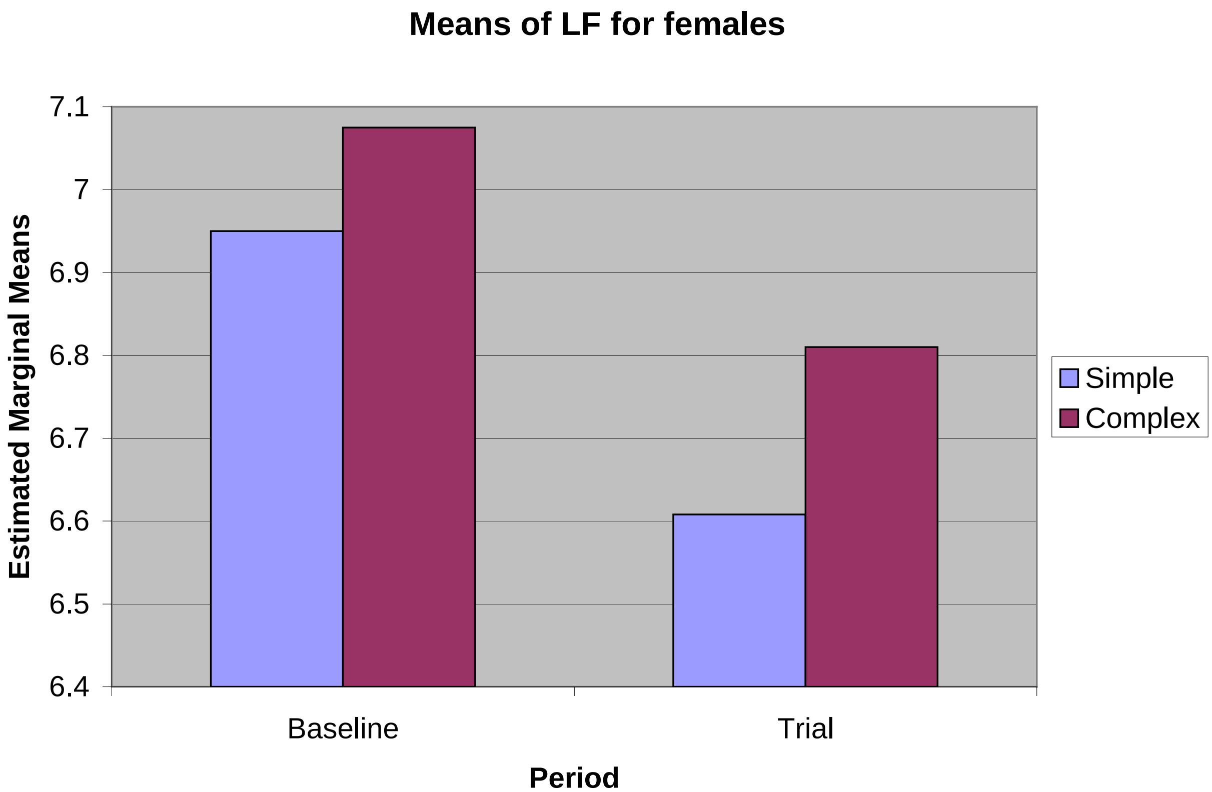

Figure 8A. Means for LF for females, collapsed across songs.

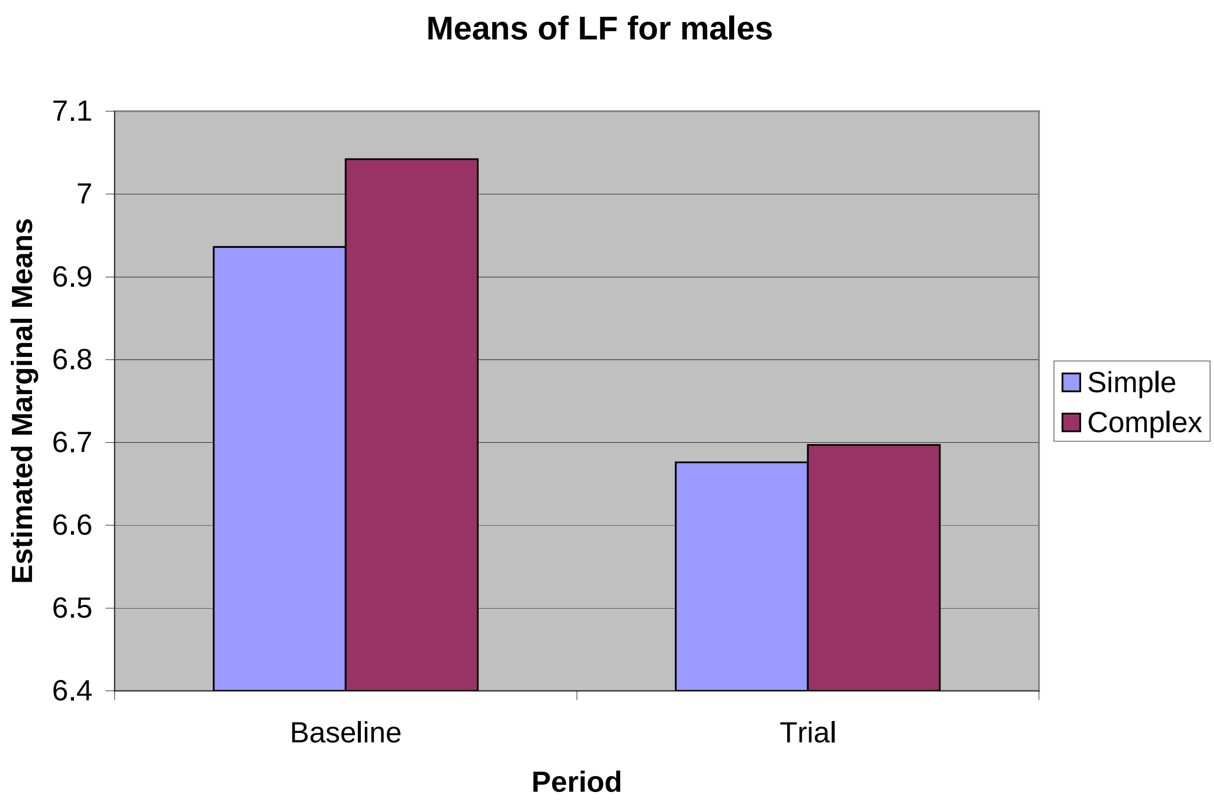

Figure 8B. Means for LF for males, collapsed across songs.

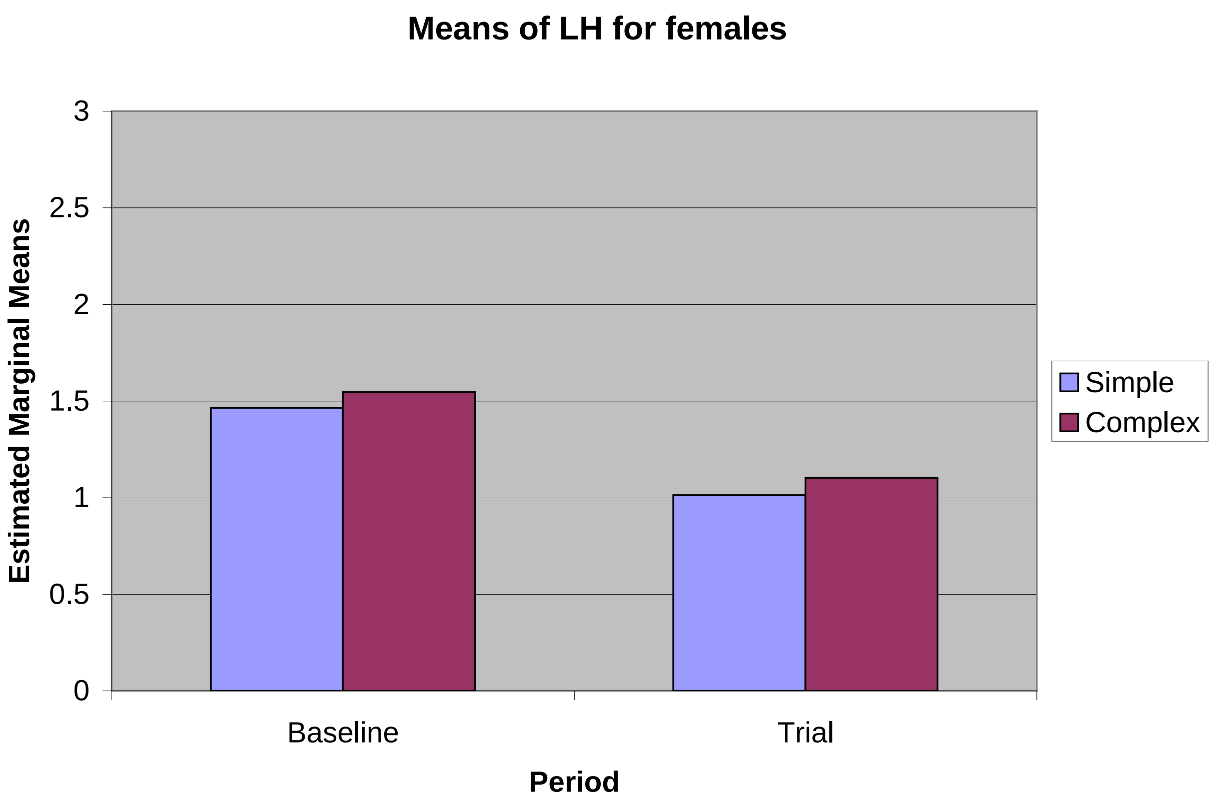

Figure 9A. Means for LH for females, collapsed across songs.

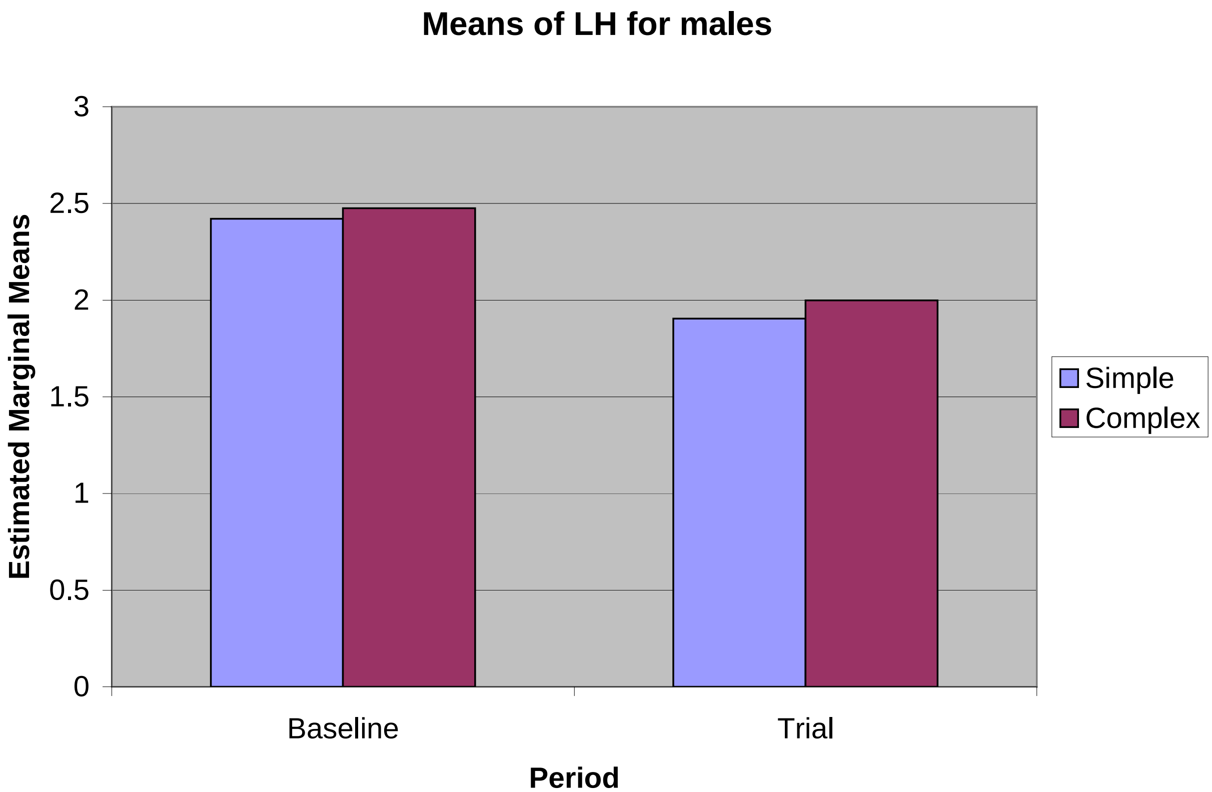

Figure 9B. Means for LH for males, collapsed across songs.

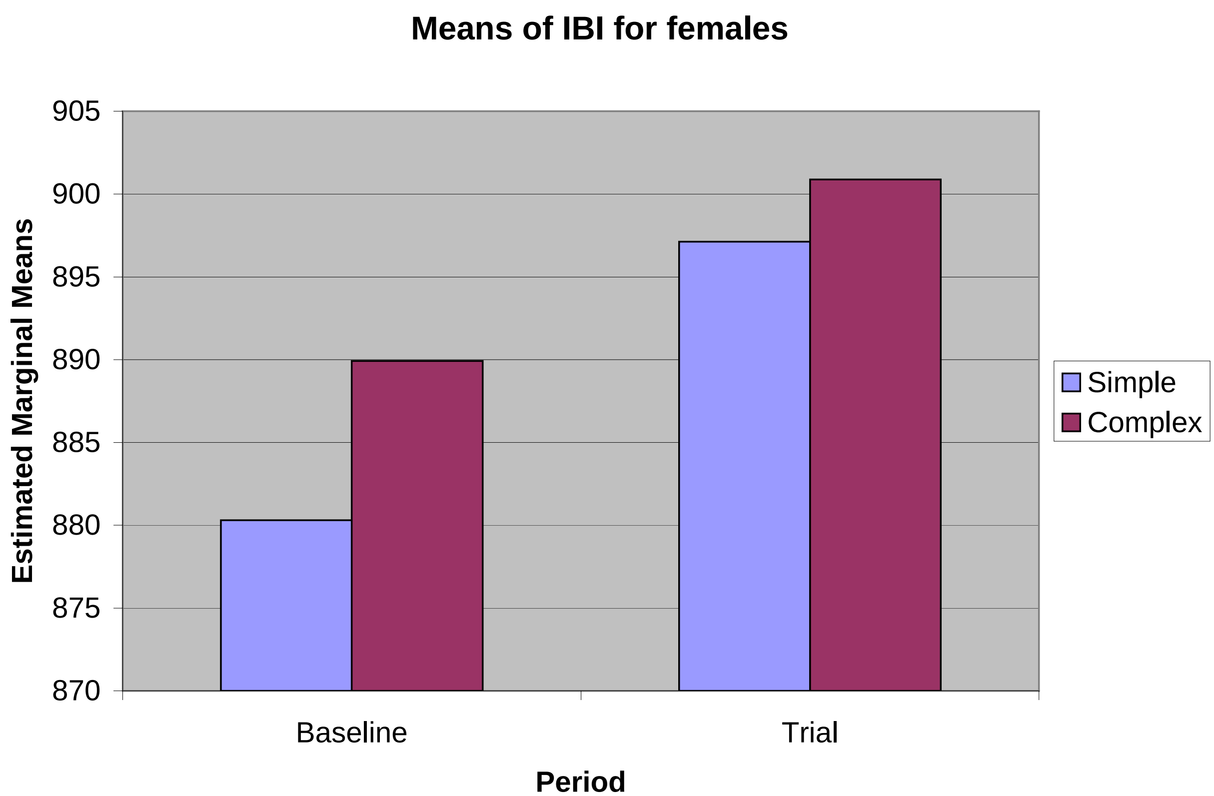

Figure 10A. Means for IBI for females, collapsed across songs.

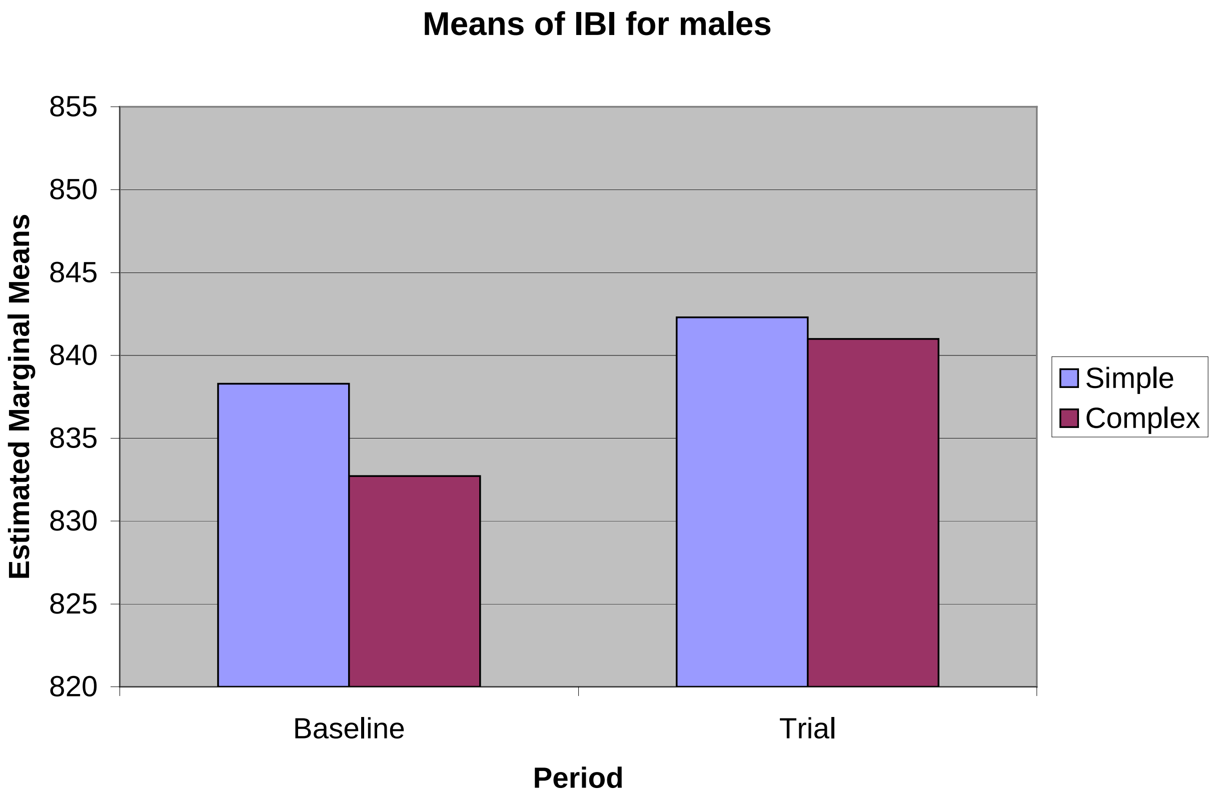

Figure 10B. Means for IBI for males, collapsed across songs.

$$\left(\left(\frac{\text{tri-ac} + \text{tri-bc} + \text{tri-cc}}{3}\right) - \left(\frac{\text{pre-ac} + \text{pre-bc} + \text{pre-cc}}{3}\right)\right) - \left(\left(\frac{\text{tri-as} + \text{tri-bs} + \text{tri-cs}}{3}\right) - \left(\frac{\text{pre-as} + \text{pre-bs} + \text{pre-cs}}{3}\right)\right)$$

Figure 11. Graphic representation of double-difference formula.

tri = trial
pre = baseline
a,b,c = song identifiers
c,s = complex, simple

Note that this example is shown with three songs, but all five songs were used in the actual calculation.

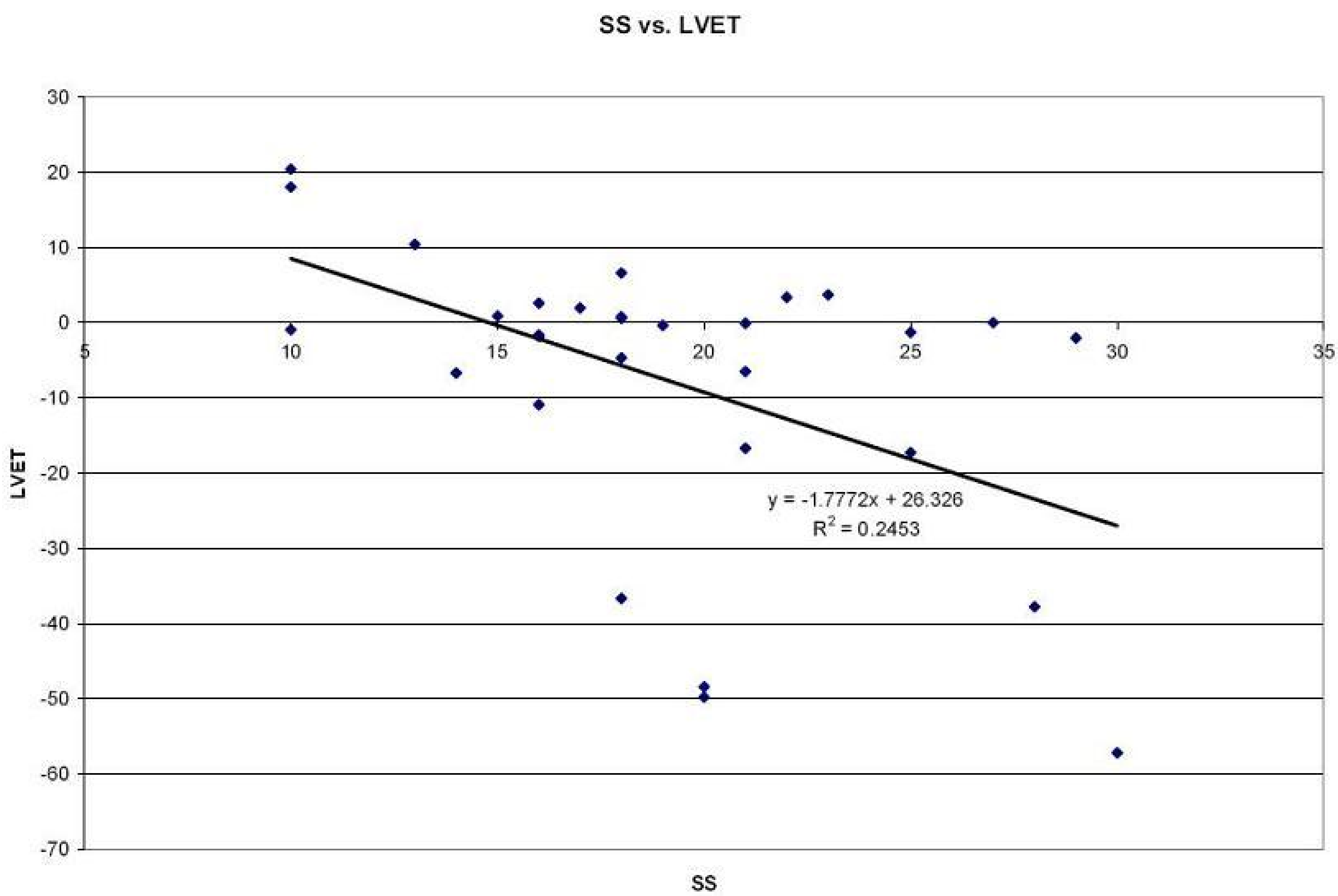

Figure 12. LVET double-difference scores plotted against SS.

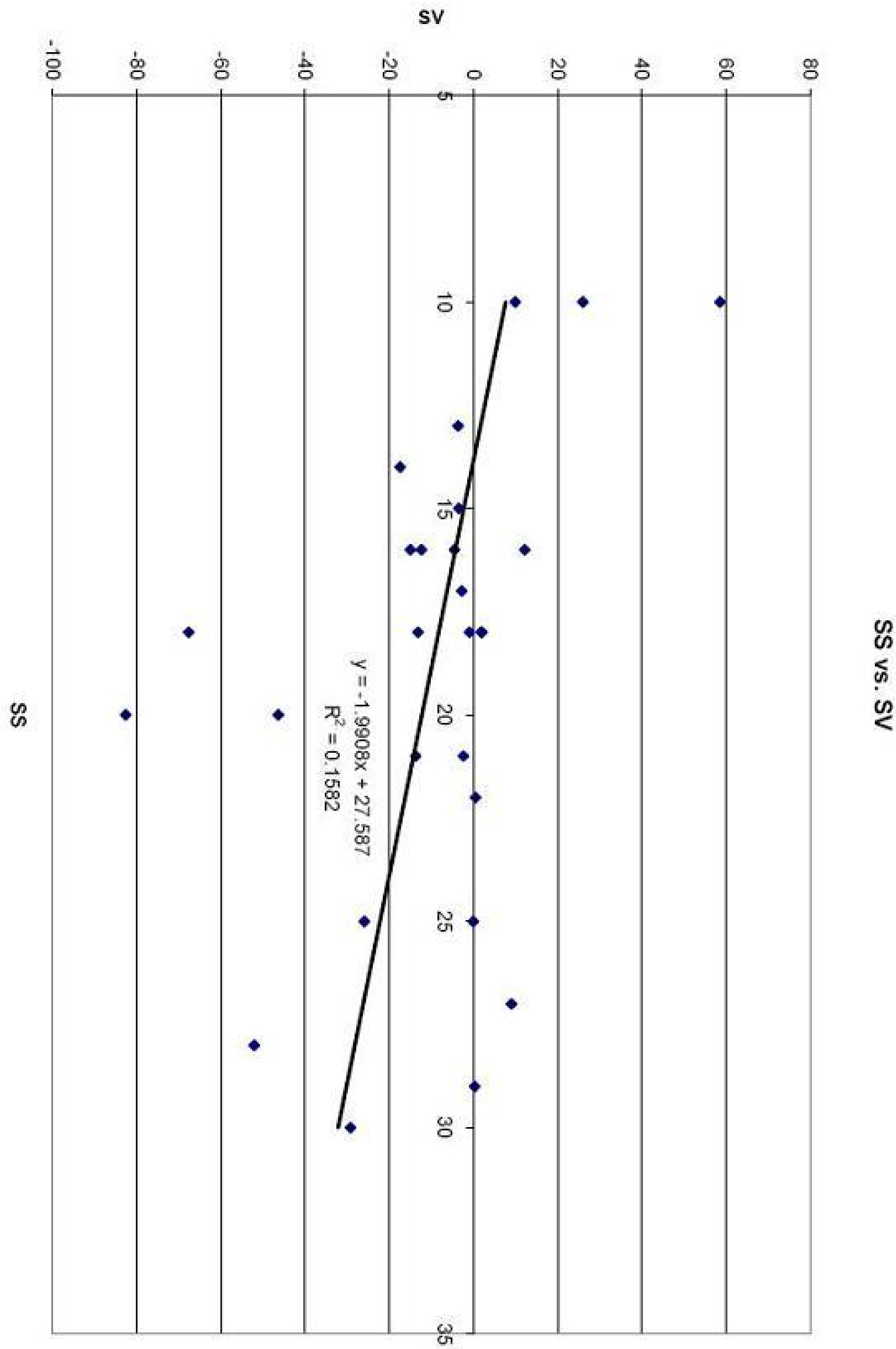

Figure 13. SV double-difference scores plotted against SS.

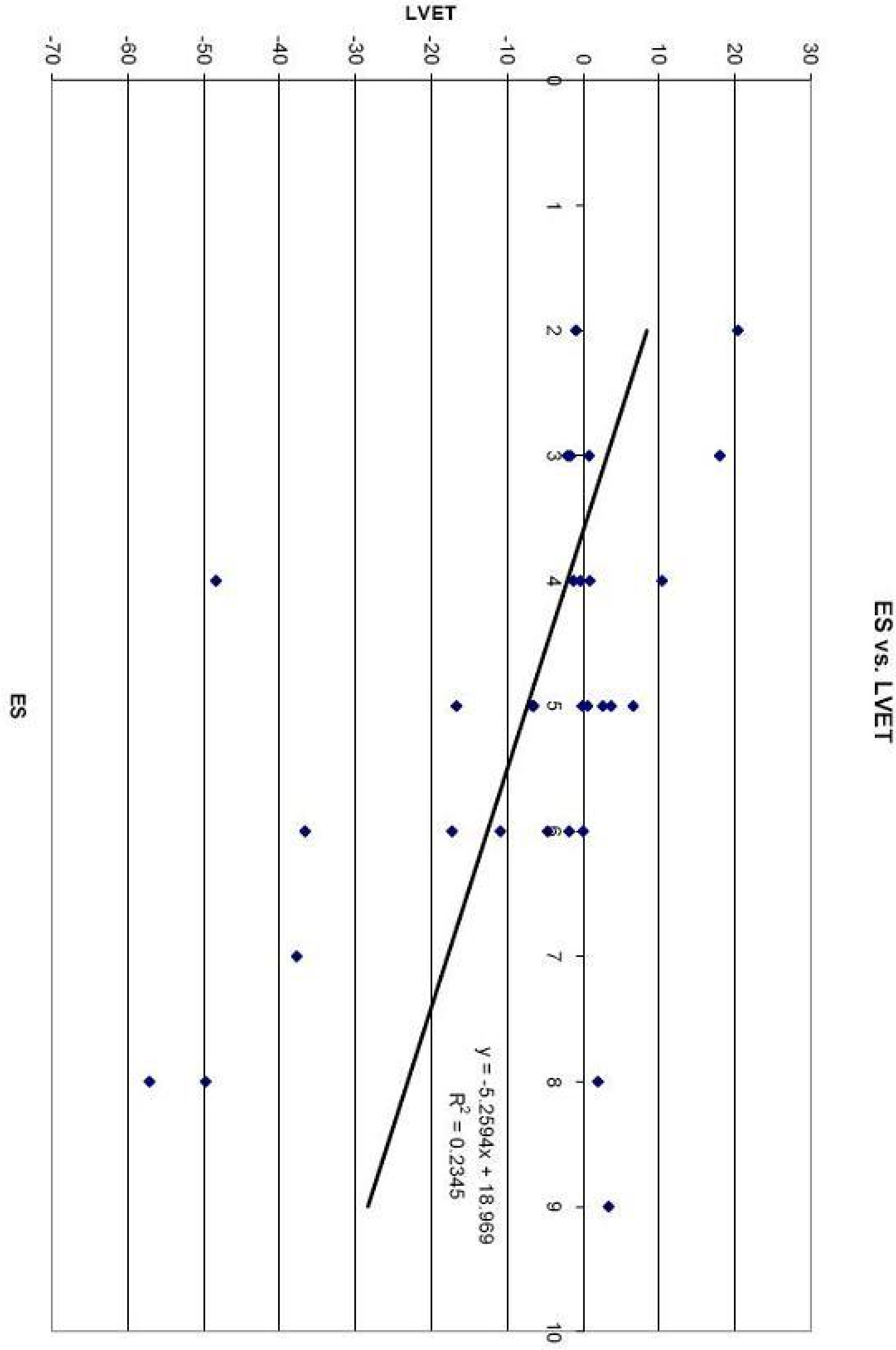

Figure 14. LVET double-difference scores plotted against ES.

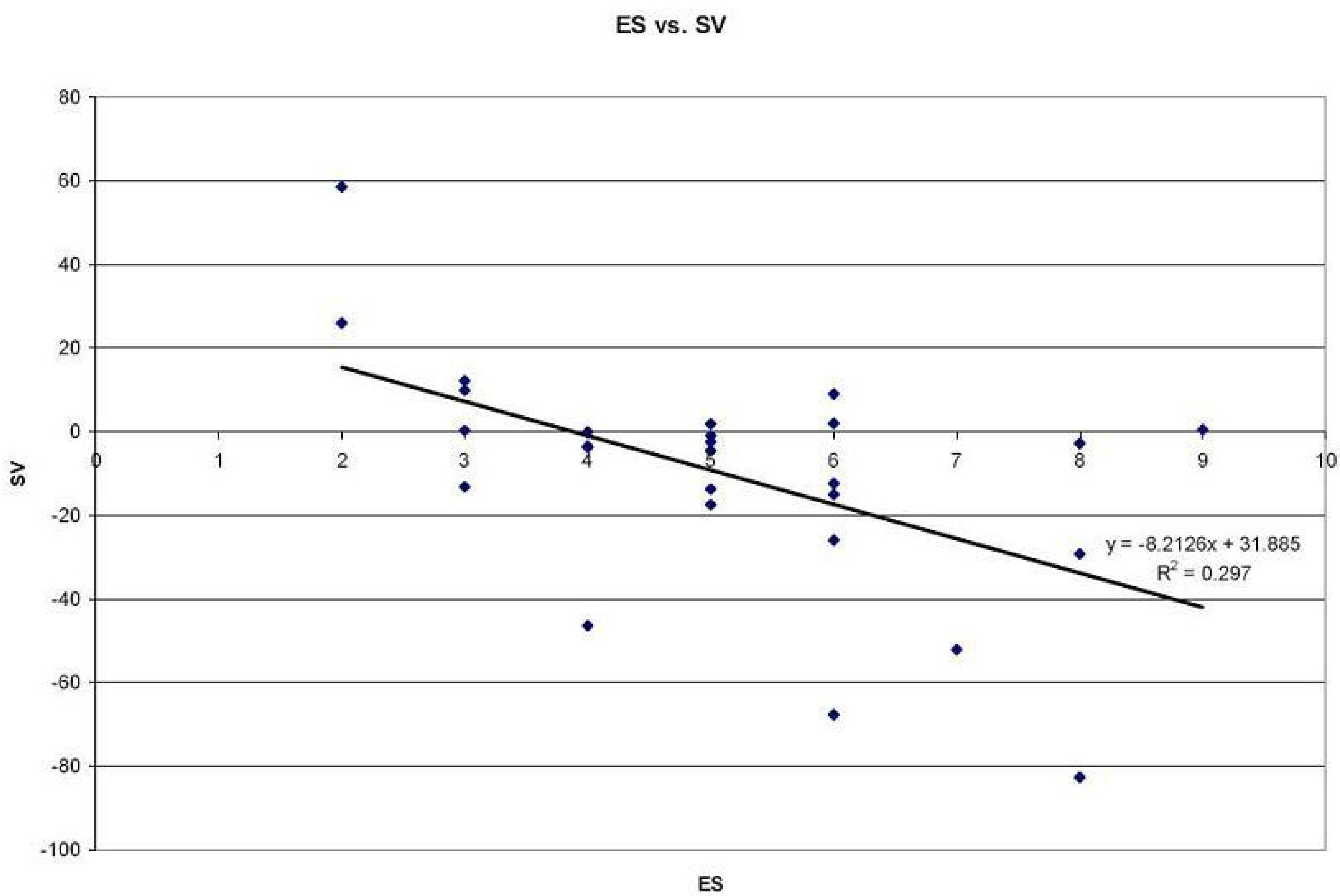

Figure 15. SV double-difference scores plotted against ES.

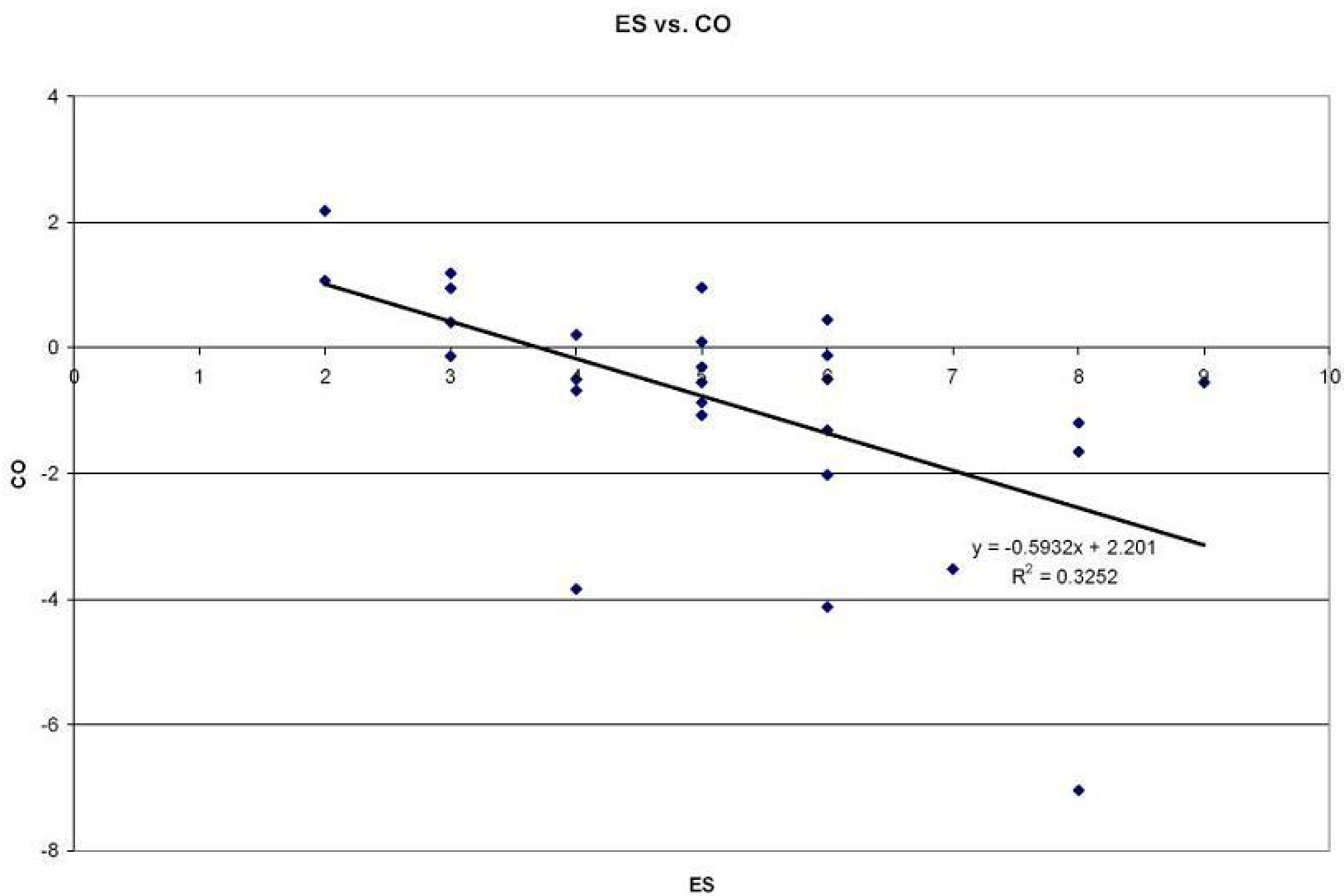

Figure 16. CO double-difference scores plotted against ES.

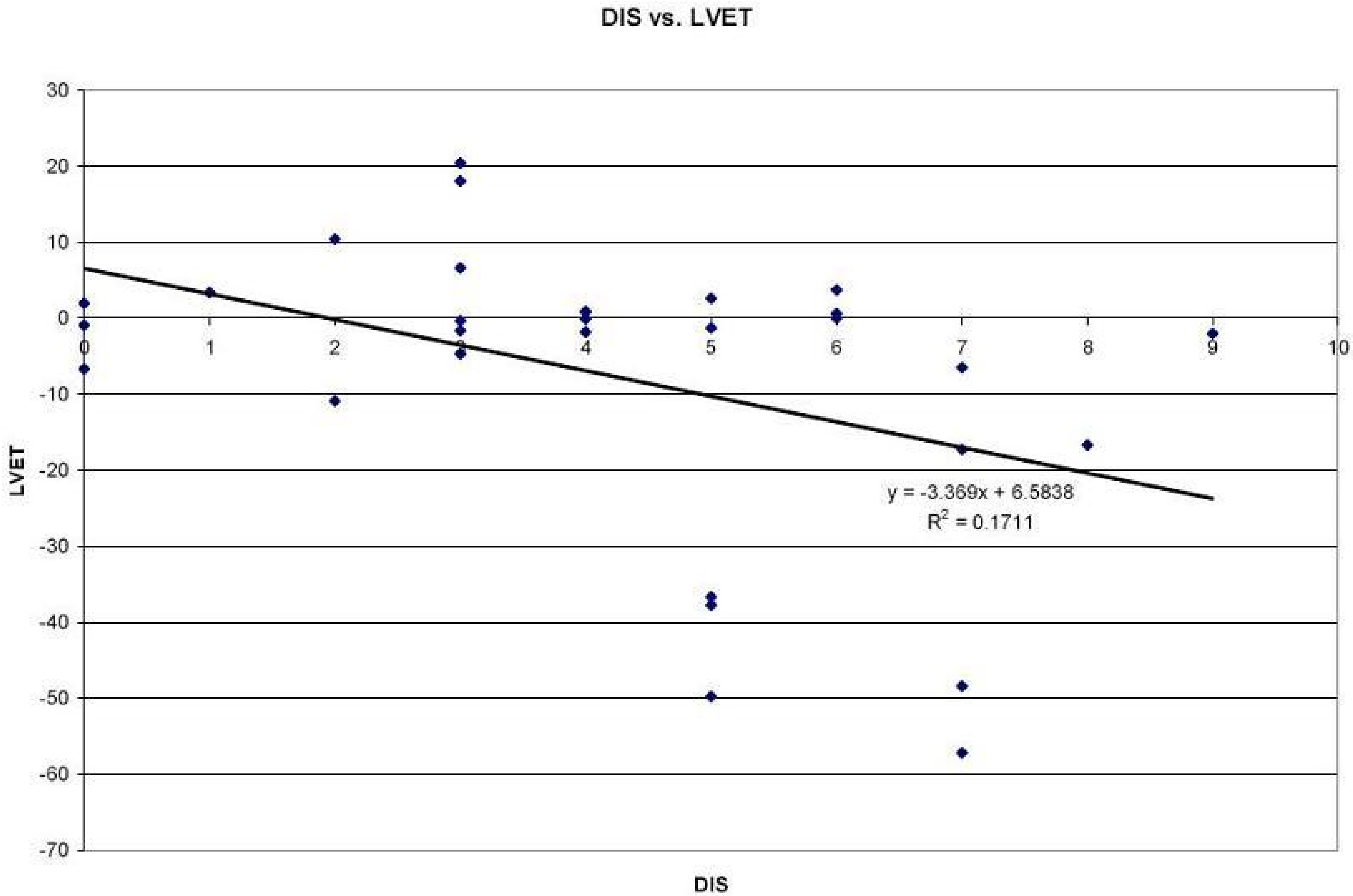

Figure 17. LVET double-difference scores plotted against DIS.

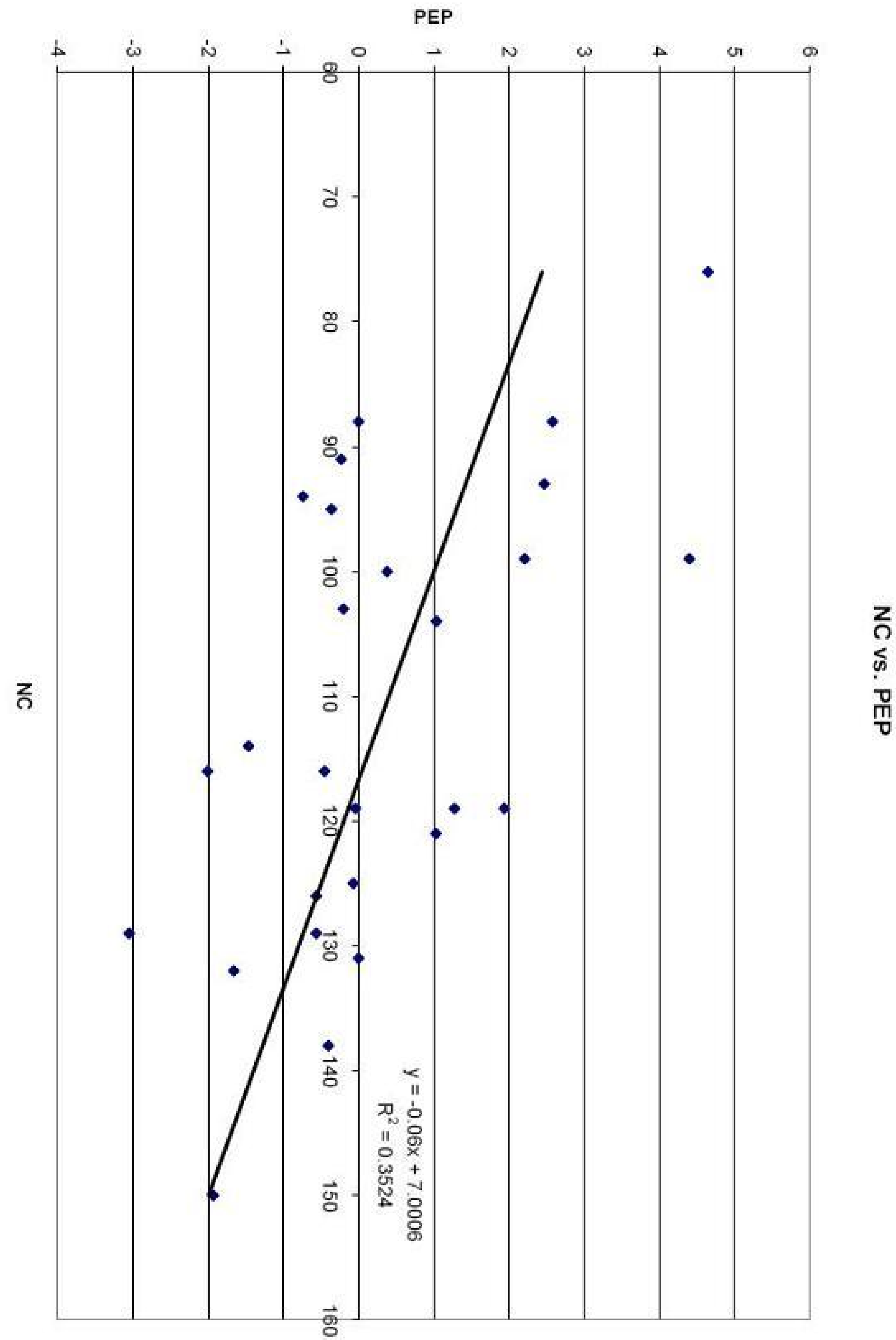

Figure 18. PEP double-difference scores plotted against NC.

Appendix A

1. Sex: A = female; B = male.
2. Do you have any current hearing deficiencies? A = no; B = yes (if yes, please explain).
3. Number of years of music theory classes and/or years of private tutoring in music theory within the past 5 years: A = 0; B = 1; C = 2; D = 3; E = more than 3.
4. Number of years of instrumental or vocal music lessons, either private or group within the past 5 years: A = 0; B = 1; C = 2; D = 3; E = more than 3.
5. Number of years as a musical performer within the past 5 years: A = 0; B = 1; C = 2; D = 3; E = more than 3.
6. Rate your own general level of understanding of music: A = I don’t understand anything about music; B = I understand very little about music; C = I understand some aspects of music; D = I understand most aspects of music; E = I understand almost all aspects of music.
7. How important has music been in your life in the past 3 years? A = not at all; B = hardly at all; C = moderately; D = very; E = extremely.
8. On average, how many hours per day do you actually spend listening to music, either while doing something else or as your main activity? A = 0; B = 1-2; C = 3-4; D = 5-8; E = 9 or more.
9. How much time would you prefer to be able to spend listening to music? A = 0; B = 1-2; C = 3-4; D = 5-8; E = 9 or more.
10. What is your usual level of involvement when you listen to music? A = background only; B = hardly at all; C = moderately; D = very; E = total concentration.
11. How many musical events (concerts, recitals, clubs, etc.) have you attended in the past 12 months? A = 0; B = 1-3; C = 4-7; D = 7-9; E = 10 or more.

Appendix B

Please answer the questions below using the following scale:

1	2	3	4	5
Not at all	Slightly	Moderately	Very	Extremely

1. How complex was the piece you just heard?
2. How exciting was the piece you just heard?
3. How much did you like the piece of music you just heard?
4. How happy was the piece you just heard?
5. How relaxing was the piece you just heard?
6. How sad was the piece you just heard?

[Note: Items 5 and 6 are reverse-scored and averaged with items 2 and 4, respectively. Questionnaires with answers that substantially disagreed were thrown out of analysis.]

References

Allen, M., & Crowell, M. (1989). Patterns of autonomic response during laboratory stressors. *Psychophysiology*, *26*, 603-614.

Bartlett, D. (1996). Physiological responses to music and sound stimuli. In D.A. Hodges (Ed.), *Handbook of music psychology* (pp. 345-385). IMR Press, San Antonio, TX.

Baumgartner, T., Esslen, M., & Jäncke, L. (2006). From emotion perception to emotion experience: Emotions evoked by pictures and classical music. *International Journal of Psychophysiology*, *60*, 34-43.

Berlyne, D. (1970). Novelty, complexity, and hedonic value. *Perception and Psychophysics*, *8*, 279-286.

Berlyne, D., Craw, M., Salapatek, P., & Lewis, J. (1963). Novelty, complexity, incongruity, extrinsic motivation, and the GSR. *Journal of Experimental Psychology*, *66*, 560-567.

Berlyne, D., McDonnell, P., Nicki, R., & Parham, L. (1967). Effects of auditory pitch and complexity on EEG desynchronization and on verbally expressed judgments. *Canadian Journal of Psychology*, *21*, 346-367.

Bernardi, L., Porta, C., & Sleight, L. (2006). Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: The importance of silence. *Heart*, *92*, 445-452.

Birbaumer, N., Lutzenberger, W., Rau, H., Mayer-Kress, G., & Braun, C. (1996). Perception of music and dimensional complexity of brain activity. *International Journal of Bifurcations and Chaos*, *6*, 267-278. Retrieved October 5, 2006, from http://www.santafe.edu/~gmk/Pubs/MusicEEG/

Burke, M., & Gridley, M. (1990). Musical preferences as a function of stimulus complexity and listeners' sophistication. *Perceptual and Motor Skills*, *71*, 687-690.

Cacioppo, J., & Petty, R. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*, 116-131.

Cacioppo, J., Petty, R., & Kao, C. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, *48*, 306-307.

Collins, R., Litman, J., & Spielberger, C. (2004). The measurement of perceptual curiosity. *Personality and Individual Differences*, *36*, 1127-1141.

Eerola, T., & North, A. (2000). Expectancy-based model of melodic complexity. In C. Woods, G. B. Luck, R. Brochard, S. A. O'Neill, and J. A. Sloboda (Eds.) *Proceedings of the sixth international conference on music perception and cognition*. Keele, UK. Retrieved October 13, 2006, from http://www.cc.jyu.fi/~ptee/publications/3_2000.pdf

Etzel, J., Johnsen, E., Dickerson, J., Tranel, D., & Adolphs, R. (2006). Cardiovascular and respiratory responses during musical mood induction. *International Journal of Psychophyiosology*, *61*, 57-69.

Furnham, A., & Allass, K. (1999). The influence of musical distraction of varying complexity on the cognitive performance of extroverts and introverts. *European Journal of Personality*, *13*, 27-38.

Furnham, A., & Bradley, A. (1997). Music while you work: The differential distraction of background music on the cognitive test performance of introverts and extraverts. *Applied Cognitive Psychology*, *11*, 445-455.

Gomez, P., & Danuser, B. (2004). Affective and physiological responses to environmental noises and music. *International Journal of Psychophysiology*, *53*, 91-103.

Haas, F., Distenfeld, S., & Axen, K. (1986). Effects of perceived musical rhythm on respiratory pattern. *Journal of Applied Physiology*, *61*, 1185-1191.

Iwanaga, M., Kobayashi, A., & Kawasaki, C. (2005). Heart rate variability with repetitive exposure to music. *Biological Psychology*, *70*, 61-66.

Iwanaga, M., & Moroki, Y. (1999). Subjective and physiological responses to music stimuli controlled over activity and preference. *The Journal of Music Therapy*, *36*, 26-38.

Iwanaga, M., & Tsukamoto, M. (1997). Effects of excitative and sedative music on subjective and physiological relaxation. *Perceptual and Motor Skills*, *85*, 287-296.

Kiger, D. (1989). Effects of music information load on a reading comprehension task. *Perceptual and Motor Skills*, *69*, 531-534.

Levy, M. (1971). Sympathetic-parasympathetic interactions in the heart. *Ciruclation Research*, *29*, 437-445.

Litle, P, & Zuckerman, M. (1986). Sensation seeking and music preferences. *Personality and Individual Differences*, *7*, 575-577.

Malliani, A. (1999). The pattern of sympathovagal balance explored in the frequency domain. *News in Physiological Sciences*, *14*, 111-117.

Marteau, T., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, *31*, 301-306.

Nater, M., Krebs, M., & Ehlert, U. (2005). Sensation seeking, music preference, and psychophysiological reactivity to music. *Musicae Scientiae*, *9*, 239-254.

Nater, M., Abbruzzese, E., Krebs, M., & Ehlert, U. (2006). Sex differences in emotional and psychophysiological responses to musical stimuli. *International Journal of Psychophysiology*, *62*, 300-308.

North, A., & Hargreaves, D. (1995). Subjective complexity, familiarity, and liking for popular music. *Psychomusicology*, *14*, 77-93.

North, A., & Hargreaves, D. (1999). Can music move people? The effects of musical complexity and silence on waiting time. *Environment and Behavior*, *31*, 136-149.

Nyklíček, I., Thayer, J., & Van Doornen, L. (1997). Cardiorespiratory differentiation of musically-induced emotions. *Journal of Psychophysiology*, *11*, 304-321.

Orr, M., & Ohlsson, S. (2005). Relationship between complexity and liking as a function of expertise. *Music Perception*, *22*, 583-611.

Parry, R. (2004). Musical complexity and Top 40 Chart performance. In *GVU technical report* [Online database], Georgia Institute of Technology, Atlanta. Retrieved October 5, 2006, from http://smartech.gatech.edu/handle/1853/50

Pressing, J. (1998). Cognitive complexity and the structure of musical patterns. In *Proceedings of the 4th conference of the Australasian cognitive science society*, Newcastle, Australia. http://www.psych.unimelb.edu.au/staff/jp/cog-music.pdf

Saul, J. (1990). Beat-to-beat variations of heart rate reflect modulation of cardiac autonomic outflow. *News in Physiological Sciences*, *5*, 32-37.

Scheirer, E., Watson, R., & Vercoe, B. (2000). On the perceived complexity of short musical segments. In C. Woods, G. B. Luck, R. Brochard, S. A. O'Neill, and J. A. Sloboda (Eds.) *Proceedings of the sixth international conference on music perception and cognition*. Keele, UK. Retrieved October 14, 2006, from http://eds.www.media.mit.edu/people/eds/icmpc00.pdf

Shmulevich, I., & Povel, D. (2000). Measures of temporal pattern complexity. *Journal of New Music Research*, *29*, 61-69.

Steck, L., & Machotka, P. (1975). Preference for musical complexity: Effects of context. *Journal of Experimental Psychology*, *104*, 170-174.

Streich, S. (2005). Automatic characterization of music complexity: a multi-faceted approach. Unpublished doctoral dissertation, Universitate Pompeu Fabra, Barcelona, Spain. Retrieved October 7, 2006 from http://www.iua.upf.edu/mtg/publications/e29495-DEA-sstreich2005.pdf

Tanaka, G., Sawada, Y., & Fujii, R. (1994). Cardiac vagal withdrawal during stress tasks estimated by the barometric sensitivity and respiratory sinus arrhythmia. *The Japanese Journal of Psychology*, *65*, 9-17.

Thayer, J. (1986). Multiple indicators of affective response to music. Unpublished doctoral dissertation, New York University.

Uijtdehaage, S., & Thayer, J. (2000). Accentuated antagonism in the control of human heart rate. *Clinical Autonomic Research*, *10*, 107-110.

Yanagihashi, R., Ohira, M., Kimura, T., & Fujiwara, T. (1997). Physiological and psychological assessment of sound. *International Journal of Biometeorology*, *40*, 157-161.

Zuckerman, M. (1990). The psychophysiology of sensation seeking. *Journal of Personality*, *58*, 313-345.

Zuckerman, M. (1994). *Behavior expressions and biosocial bases of sensation seeking*. New York: Cambridge University Press.